The accumulation of deleterious mutations
by Muller's ratchet

Isabel Mendes Gordo

PhD Thesis
2001
Ó mar salgado, quanto do teu sal
São lágrimas de Portugal!
Por te cruzarmos, quantas mães choraram,
Quantos filhos em vão rezaram!

Quantas noivas ficaram por casar
Para que fosses nosso, ó mar!
Valeu a pena? Tudo vale a pena
Se a alma não é pequena.

Quem quer passar além do Bojador
Tem que passar além da dor.
Deus ao mar o perigo e o abismo deu,
Mas nele é que espelhou o céu.

Fernando Pessoa
Acknowledgments

I would like to express my many thanks to the following people, without whom I could not have written this thesis.

My supervisor Brian Charlesworth, who first taught me the basic principles of population genetics, for having opened my horizon to the fantastic world of evolutionary genetics and having guided me throughout this study with infinite patience and constant encouragement. My second supervisor Peter Keightley for his suggestions, guidance and helpful discussions. Both were extremely helpful with the manuscripts.

Deborah Charlesworth and Gil McVean for help with the simulation work, numerous comments on manuscripts and posters, and for constant discussions. Carolina Bartolome, Frantz Depaulis, Valerie Laporte, Xulio Maside, Arcadi Navarro and Jorge and Cristina Vieira for having answered all the questions I have made and helping me many times. Peter Andolfatto and Stephen Wright both for their comments and discussions and also for being excellent office colleagues. Doris Bachtrog for her friendship, support, countless discussions and helpful advice.

My parents for having taught me that knowledge is priceless and made me believe that I could look for it. My brother Paulo, who offered me a book on biology while wishing that I would continue studying physics, my brother Ze Manuel, and my friend Ze Pinheiro for all their encouragement.

My husband Manuel for having understood my wishes, reminded me about them when I would forget, and having always stayed by my side.

I dedicate this thesis to my two nieces and two nephews, with whom I share the annoying tendency of asking why.
Abstract

The evolutionary significance of genetic recombination is one of the most intriguing problems in evolutionary biology, since recombination is one of the primary features of sexual reproduction. Of the vast number of questions that one can ask in relation to sex and recombination, the simplest is: what happens in their absence? It was realised long ago that the level of recombination influences the action of natural selection. Non-recombining genomes are expected to adapt more slowly than recombining ones. In addition, they are also more prone to degeneration by the accumulation of harmful mutations. This study is concerned with the latter process: the fate of a non-recombining genome or chromosome that is continuously subject to recurrent mutation to deleterious alleles.

H.J. Muller argued that one major difference between a non-recombining asexual and a recombining sexual population is that, in the former, genetic drift can overwhelm selection against deleterious mutations, whereas this is not so likely in the presence of recombination. In the absence of recombination, deleterious mutations can therefore accumulate by what Muller called a ratchet-like process. This study focuses on four aspects of this mechanism. First, the quantification of its rate is examined, both by simulation methods and by analytical approximations, under the simplest possible model. Second, its interaction with another phenomenon, the continuous elimination of strongly deleterious alleles, is studied. Third, its effects on neutral DNA polymorphism, and the possibility of detecting its action by measuring these effects, are studied. Fourth, the circumstances under which this process may have significance for the evolution of non-recombining Y chromosomes are analysed.

It is hoped that this work represents a useful contribution to a better understanding of Muller's ratchet. The overall conclusion is that, given the present
empirical knowledge on rates and effects of deleterious mutations and on levels and patterns of variability on Y chromosomes, Muller’s ratchet may well be significant in driving their evolution, even in species with large populations.
Contents

1 Introduction

1.1 Selective Sweeps
1.2 Background Selection
1.3 Muller’s ratchet
  1.3.1 Evolutionary advantage of recombination
  1.3.2 Extinction of asexual populations
  1.3.3 The degeneration of the Y chromosome
  1.3.4 Selfing populations
  1.3.5 What stops or slows Muller’s ratchet?
1.4 Weak Hill-Robertson effects
1.5 Relating models to observations

2 The speed of Muller’s ratchet

2.1 Introduction
2.2 Approximation based on diffusion equation
2.3 Simulation methods
2.4 Simulation Results
2.5 The time between clicks and the decline in mean fitness
2.6 Discussion

3 The speed of Muller’s ratchet with Background Selection, and
the degeneration of Y chromosomes

3.1 Introduction
3.2 Theoretical considerations
  3.2.1 Muller’s ratchet and Background Selection
3.3 Simulation methods
3.4 Results

3.4.1 The rate of accumulation by the ratchet with background selection

3.4.1.1 Region of parameter space for which $n_{oo} > 1$

3.4.1.2 Region of parameter space for which $n_{oo} < 1$

3.4.2 Dynamics of very weak deleterious mutations

3.4.3 Accumulation of both types of mutations

3.5 Discussion

3.5.1 The rate of the ratchet under background selection

3.5.2 The rate of the ratchet in small populations

3.5.3 Y and neo-Y chromosomes

3.5.4 The accumulation of transposable elements due to Muller’s ratchet

4 The level and pattern of neutral variation under Muller’s ratchet

4.1 Introduction

4.2 Simulation methods

4.2.1 Multi-locus Monte Carlo simulations

4.2.2 Coalescents simulations

4.2.3 Measures of genetic diversity at a neutral locus

4.2.4 The mutational frequency spectrum

4.3 Results

4.3.1 Muller’s ratchet and genetic diversity

4.3.2 Muller’s ratchet and frequency spectrum

4.4 Muller’s ratchet with two types of deleterious mutations

4.5 Discussion

4.5.1 Muller’s ratchet and neutral variation
4.5.2 Muller's ratchet and the Y chromosome  
4.5.3 Muller's ratchet and a non-recombining autosomal region

5 Muller's ratchet and the distribution of selection coefficients

5.1 Introduction  
5.2 Simulation methods  
5.2.1. Monte Carlo simulation for the speed  
5.2.2. Monte Carlo simulations for neutral variability  
5.3 Preliminary considerations  
5.4 Results  
5.4.1 The ratchet and the decline in mean fitness  
5.4.2 The ratchet and the level of diversity  
5.5 Discussion  
5.5.1 Speed of Muller's ratchet  
5.5.2 Muller's ratchet and neutral variation  
5.5.3 Muller's ratchet and the Y chromosome

Bibliography
1 Introduction

This chapter is based on a review article published in *Current Biology* on the evolutionary consequences of genetic linkage (Gordo and Charlesworth 2001a).

The rate of genetic recombination varies among species, and across different regions of the genome within a species. In the most extreme case, all genes in an asexual or self-fertilizing species are effectively completely linked, in contrast to the substantial opportunities for recombination in outbreeding sexual species. Since two of the major model organisms, *Caenorhabditis elegans* and *Arabidopsis thaliana*, are highly self-fertilizing hermaphrodites, the evolutionary consequences of restricted recombination should be of interest to a wide spectrum of biologists. Here I briefly review both empirical findings and theoretical models that deal with the general question of the relevance of recombination in evolutionary progress.

Several interesting observations on the relation between patterns of evolution and levels of genetic recombination have already been made.

First, asexual or highly self-fertilizing species tend to be young, in terms of the evolutionary timescale, suggesting that they become extinct more rapidly than their recombining relatives (Maynard Smith 1978; Stanley 1975; Takebayashi and Morrell 2001). Second, organelle and clonally transmitted genomes in several taxa show reduced levels of adaptation with respect to RNA and protein sequences (Lynch 1996; Lynch 1997; Nachman 1998). The population genetic processes discussed on this thesis have been proposed by several authors to be as a contributing factor to this observation.

Second, in *Drosophila*, genomic regions where recombination is reduced in frequency (such as the centromeric and telomeric regions) show reduced levels of codon bias (the non-random usage of alternative codons corresponding to the same amino acid), suggesting a reduction in the ability of selection to maintain this aspect.
of molecular adaptation (Comeron et al. 1999; Kliman and Hey 1994; Powell and Moriyama 1997).

Finally, in a number of different species it has been found that genomic regions with restricted recombination show lower levels of DNA sequence variation than more highly recombining regions (Andolfatto and Przeworski 2001; Begun and Aquadro 1992; Nachman et al. 1998; Stephan and Langley 1998; Stephan et al. 1998; Zurovcova and Eanes 1999). One of the most spectacular examples of these patterns is provided by non-recombining Y chromosomes, or neo-Y chromosomes formed by fusions between autosomes and sex chromosomes, for which there is evidence of both degeneration of gene function and reduced genetic variability (Bachtrog and Charlesworth 2000; Charlesworth and Charlesworth 2000; Yi 2000).

In this thesis, I will discuss the evolution of Y and neo-Y chromosomes in the context of processes that can lead to both loss of their gene function and reduction of their DNA sequence variation.

Why should there be a relation between the amount of recombination and levels of variation and adaptation? As Fisher and Muller pointed out long ago (Fisher 1930; Muller 1964), the dynamics of a given gene is not only influenced by the evolutionary forces acting on the gene itself, but also by forces acting at linked loci. This means that the predictions of single-locus population genetics have to be modified when selection is acting on sets of linked loci. In the following sections, some of the processes that have been proposed as important in shaping evolution when recombination is restricted over a sizeable genomic region, will be described. They provide some possible explanations for the observations just mentioned. All of these processes, and in particular the one with which this study is concerned, reflect a general effect first quantified by Hill and Robertson: a locus linked to another locus under directional selection experiences a reduced effective population size, $N_e$ (Felsenstein 1974; Hill and Robertson 1966). The extent of random fluctuations in
allele frequencies due to finite population size, genetic drift, is inversely related to $N_e$, so that this effect means that selective differences at one locus tend to enhance the effects of drift at a linked locus.

To understand the role of the Hill-Robertson effect in molecular evolution and variation, it is necessary to review some basic single-locus population genetics results. According to the neutral theory of molecular evolution, mutation and genetic drift have a major influence on DNA sequence variation within species and on differences between species. Mutation creates new neutral variants (with no significant fitness effects), and genetic drift causes random changes in their frequencies until fixation or loss. Kimura showed that the level of variation within a population at a neutral locus is proportional to the product of $N_e$ and the neutral mutation rate $\mu$, and that the rate of sequence evolution is equal to $\mu$ (Kimura 1983). A mutation with a selection coefficient of $s$ (which measures the reduction or increase in the fitness of its carriers, relative to that of the rest of the population) is effectively neutral if $N_es << 1$ (Kimura 1983). In particular, a strongly deleterious mutation (for which $N_es >> 1$) will be rapidly eliminated with high probability, but a weakly deleterious mutation (with $N_es < 1$) can survive for a long time, and even become fixed in the population, as a result of genetic drift. Similarly, a favourable mutation will have almost the same chance of loss from the population as a neutral mutation if $N_es < 1$. Changes in $N_e$ caused by different kinds of selection at linked loci can thus greatly affect both genetic variability and the efficacy of selection.

I will now describe several kinds of selection models. These will introduce the process with which this thesis is concerned, and place it in a general context.
1.1 Selective sweeps

Let us first consider the effect of an advantageous mutation on the level of variation at a completely linked, neutral locus (Figure 1.1). A peculiar footprint is left shortly after such a mutation has swept through the population to fixation: variation is drastically reduced, and any variants that can be observed are present at low frequencies in the population. From the point of view of the neutral locus, it is as if the population went through a bottleneck of one individual, and then expanded to its normal size.

![Figure 1.1](image)

**Figure 1.1** The effect of a selective sweep on neutral variation
A) In a population with a considerable amount of standing variation (black circles) at a neutral locus (white), an advantageous mutation arises at a completely linked locus (box with squares). B) Because of its higher fitness, it will increase in frequency in the population. As it spreads through the population, it wipes out variation at the neutral locus. At the time of fixation of the advantageous mutation, all the variation at the locus is lost. C) Some time after fixation, neutral variation starts to build up in the population, as a result of the occurrence of new mutations (x). Since the new neutral variants are relatively young, their frequencies in the population are very low. If we take a sample from the population and sequence alleles of the neutral locus, we will tend to observe mutations that appear only once in the sample. This pattern is very different from that expected in a population at statistical equilibrium under mutation and genetic drift, where some variants will occur at intermediate frequencies.

If an advantageous mutation occurs in a region where there is a small amount of recombination, neutral variation will not be completely lost, but will still be reduced. The level of variation in a particular region of a chromosome depends on
the rate of adaptive evolution, and on the ratio of the strength of selection to the recombination rate in that region. If selective sweeps are continuously occurring throughout the genome of a sexual species, there will be a correlation between neutral variability and the local recombination rate (Kaplan et al. 1989; Stephan et al. 1992; Wiehe and Stephan 1993).

Recurrent selective sweeps will also cause very low levels of variation in an asexual species, even if it has a large population size (Maynard Smith and Haigh 1974). In bacterial populations, sudden changes in frequencies of neutral markers has been interpreted as the signature of sweeps (Atwood et al. 1951; Guttman and Dykhuizen 1994; Imhof and Schloetterer 2001). Sweeps can also result in the fixation of linked deleterious alleles, and this has been proposed as a possible factor in the degeneration of Y chromosomes (Rice 1987). But the beneficial effect associated with the mutation that causes such phenomena has to be sufficiently large to overcome the deleterious effects of the mutations that hitchhike with it. Currently little is known about the values of the selective advantage associated with newly arisen beneficial mutations (but see Imhof and Schloetterer (2001) for some estimates for bacterial populations).

1.2 Background Selection

Any natural population is subject to a continual rain of deleterious mutations at loci throughout the genome. If such mutations are strongly selected against, they are constantly being eliminated. Such purifying selection can have a remarkable effect when linkage is complete (Figure 1.2).

Because every chromosome with a deleterious mutation is destined to be lost sooner or later, the ancestry of future generations is derived from the fraction $f_0$ of the current population that is free of deleterious mutations (Fisher, 1930 pg. 122).
The population thus has its effective size and level of neutral variation reduced by a factor of $f_0$ (Charlesworth et al. 1993a). The rate of neutral evolution is independent of $N_e$, and so remains unchanged (Charlesworth 1994; Charlesworth et al. 1993a).

**Figure 1.2** The effect of background selection

(A) Consider a non-recombining population of size $N_e$, with $m$ mutable loci. Each locus undergoes recurrent mutations to deleterious alleles (x), at a rate $u$ per locus per generation. If the deleterious effect ($s$) of a mutation is such that $N_es >> 1$, deleterious alleles stay at low frequencies, and back-mutation from mutant to wild-type alleles is negligible. In such a population, a balance between the rates of production of new mutations and of their elimination by natural selection is established. The population is divided into various classes: the *least-loaded class*, free of deleterious mutations (chromosomes in the gray zone), and other classes having some deleterious mutations (in the white zone, chromosomes have 1, 2 etc. mutations). All the chromosomes with deleterious mutations are eventually eliminated by selection, and new ones are derived from chromosomes free of such mutations (as indicated by black arrows). The least-loaded class is the ultimate source of all future lineages in the population. With independent effects of different loci on fitness, its frequency is $f_0=\exp(-U/s)$, where $U=mu$ is the deleterious mutation rate for the entire chromosome. Suppose now that a neutral or weakly selected variant arises (represented as a circle). If this variant is on a chromosome with one or more strongly deleterious mutations, it will soon be eliminated. The only way it can survive is if it arises in the least-loaded class— this effect is known as *background selection* (B).
Now consider the effect of background selection when there is a certain amount of recombination \((r)\), between a neutral variant and a deleterious mutation with which it is associated. The neutral variant can now unhitch itself from the deleterious mutation through recombination. Since the mean time spent by a deleterious allele in a large population is roughly \(1/s\), the neutral variant has a chance of survival if it unhitches itself during this time. It follows that \(r/s\) is an important quantity in determining the effect of background selection. Mathematical analysis shows that this effect can be indeed approximated by a reduction in \(N_e\) that involves \(r/s\), and yields an expression relating the level of variation in a region of the genome to the local rate of recombination (Hudson and Kaplan 1995a; Hudson and Kaplan 1995b; Nordborg et al. 1996). This can largely explain the Drosophila data (Charlesworth 1996a).

Suppose now that the sites concerned are not evolving neutrally, but are very weakly selected (as with synonymous changes to codons). It turns out that the effect of background selection can be quantified as though \(N_e\) is reduced in the same manner as for neutral sites (Charlesworth 1994; Stephan et al. 1999). Because the rate of evolution at weakly selected sites is dependent on \(N_e\), the rate of fixation of weakly deleterious mutations is increased, and that of advantageous mutations is decreased, if background selection is operating. In regions of the genome where linkage is very tight, such as the Y chromosome, levels of variability and adaptation should both be reduced. Because background selection impairs the ability of an asexual species to adapt, it may also accelerate the extinction of asexual populations (Charlesworth 1994; Peck 1994).

### 1.3 Muller's ratchet

The background selection model assumes that the frequency of the least-loaded class is stable over large periods of evolutionary time. But this need not be the case,
especially in the absence of recombination. Imagine that $f_0$ is small and that there is no recombination. Muller (1964) pointed out that genetic drift could then cause the loss of the least-loaded class in a irreversible manner—what is now classically referred to as a “click” of the ratchet (Figure 1.3). After this click, the class with one deleterious mutation becomes the new least-loaded class. But fluctuations due to sampling error in the abundance of this class can also cause it to be lost in the same way as the previous one, so that successive clicks occur, leading to a continuous accumulation of deleterious mutations to the detriment of the population’s mean fitness. This is the process now known as Muller’s ratchet (Felsenstein 1974) and its dynamics and evolutionary effects are the focus of this thesis.

![Figure 1.3. The clicking of Muller's ratchet](image)

The equilibrium distribution of the mutational classes at mutation-selection balance in a non-recombining population is shown in white, for which the population mean fitness ($W$) is simply $e^{W}$. As the size of the least-loaded class is relatively small this class is vulnerable to stochastic loss. When this happens, a “click” of the ratchet has occurred (gray). $W$ is then reduced by $(1-s)$ and the new least-loaded class now has one mutation. But this class may also be lost in the same way as the previous one, so that successive clicks occur, leading to a continuous accumulation of deleterious mutations and decline in mean fitness (dark).
Since its definition this process has been a subject of numerous investigations, both from theoreticians and empiricists, due to its potential importance in relation to various evolutionary questions, to which I am going to refer below.

1.3.1 Evolutionary advantage of recombination

As first pointed out by Muller and later by several other authors, Muller’s ratchet can create an advantage to recombination (Barton and Charlesworth 1998; Felsenstein 1974; Maynard Smith 1978; Muller 1964). Felsenstein (1974), in a review of models for the evolutionary advantage of recombination, concludes that genetic drift (the finiteness of populations) is crucial to understand the evolution of sex, because drift creates linkage disequilibrium that can be broken down by recombination. Being the first to conduct simulations on the ratchet mechanism, Felsenstein proposed that the Fisher-Muller argument (Fisher 1930; Muller 1932), which essentially states that two favourable mutations can be combined in the same individual more quickly in the presence of recombination than in its absence, together with the operation of Muller’s ratchet, could provide an intrinsic theory for the origin and persistence of recombination. The ratchet has, for a long time, been considered by several authors as a potential mechanism involved in providing some advantage to recombination, and it is indeed extensively referred as one of the major possible reasons for the fact that genetic recombination is a wide spread phenomenon (Antezana and Hudson 1997; Barton and Charlesworth 1998; Bell 1988; Crow 1999; Gessler and Xu 1999; Maynard Smith 1978; Rice 1999; West et al. 1999). Although every population is finite and suffers from recurrent deleterious mutations, the dramatic dependence of speed of the ratchet on population size (see for example Haigh (1978) and chapters 2 and 3) has led to the argument that only small asexual populations effectively suffer from its consequences. Furthermore, computer
simulations have shown that a very small amount of recombination is sufficient to stop Muller's ratchet (Charlesworth et al. 1993b) leading to the argument that the effects of this process are too weak to be involved in the evolution of the high levels of recombination (Barton and Charlesworth 1998). But it is still, in the long term, regarded as providing some advantage for sex.

In Muller's original definition of the ratchet, genetic drift is the important force for its operation. Recently however, Gessler (1995) showed that, in addition to genetic drift, mutation can actually be an important force in driving the ratchet, especially if the distribution of selection coefficients is very skewed towards mutations with weak deleterious effects. The conditions for this to be the case imply that selection is too weak to counter mutation pressure, so that the equilibrium between mutation and selection is in a state of what Gessler called a "chronic instability" (Gessler and Xu 1999). Quantitatively this will happen whenever the deterministic equilibrium size of the least-loaded class, $n_0$, is less than 1, in the absence of epistatic interactions among the loci affecting fitness. This condition can be intuitively understood, by noting that the deterministic forces of mutation and selection are of such a magnitude, that they lead the population to a very small frequency of the fittest class, such that less that one individual is expected in this class. This means that these forces lead to its loss. In these circumstances, only recombination can restore the equilibrium in the population and can therefore be selected for. Gessler (1999) showed that, in populations where a lack of equilibrium exists (due to a mutation-driven ratchet), a gene for recombination can spread to fixation. As the recombination allele invades the population, the ratchet is stopped and the population mean fitness no longer decays.
1.3.2. Extinction of asexual populations

The accumulation of mutations resulting from the action of the ratchet in small asexual populations and the consequent decline of its absolute mean fitness may result in a decrease in the population size. A smaller population size will in turn lead to increased genetic drift, and to a bigger susceptibility to the loss of the new least-loaded class. This self-accelerating ratchet is known as mutational meltdown (Gabriel et al. 1993; Lynch et al. 1993; Lynch et al. 1995; Lynch and Gabriel 1990). The positive feedback, where the rate of mutation accumulation increases with the mutational load, will eventually lead to population extinction. Gabriel et al. (1993) showed that the average time to extinction is minimised for intermediate selection coefficients. The reason is as follows: mutations with very small effects accumulate rapidly, but do not cause great damage; on the other hand, mutations with very big effects accumulate very slowly (if at all), so the highest risk of extinction is produced by mutations with slightly deleterious effects that may accumulate in a reasonable period of time and cause a relatively high decline in fitness. As a great number of sexual organisms are hosts of asexual lineages, such as mitochondria and chloroplasts, it is sometimes argued that the ratchet may be ubiquitous because it may operate in organelle genomes (Lynch 1996). However, due to the small size of organelle genomes, the genomic mutation rate is small (of the order of $10^{-4}$ per genome per year (Lynch and Blanchard 1998)), which probably makes them immune to the process.
13.3. The degeneration of the Y chromosome

The first step in the evolution of distinct sex chromosomes is attributed to the complete or partial restriction of recombination between what were once the homologous proto-X and proto-Y chromosomes. Once the exchange between proto-X and Y is shut down the proto-Y will behave like a haploid asexual population (the proto-X continues to recombine in females), and therefore may be subject to Muller’s ratchet. This was first pointed out by Charlesworth (1978), who proposed this mechanism as a general process to explain the genetic inertness of Y chromosomes or large non-recombining segments of the genome. The role of the ratchet in Y chromosome degeneration will be extensively discussed in the following chapters. Until very recently Muller’s ratchet was thought to cause a gradual increase in the number of mutant loci per chromosome with the frequency of each mutant locus being kept very low (Charlesworth 1978; Charlesworth 1996b; Rice 1987). However it has now became clear that the ratchet can indirectly cause the fixation of deleterious alleles in the whole population (Bergstrom and Pritchard 1998; Charlesworth and Charlesworth 1997; Higgs and Woodcock 1995). This fact has some relevance concerning the role of the ratchet in the evolution of the Y chromosome and the related phenomenon of dosage compensation (Charlesworth and Charlesworth 1997). Imagine that the ratchet is turning at a reasonable evolutionary speed both in terms of time between clicks as well as in the decline in fitness per click. The circumstances under which this is expected to occur will be the subject of chapter 2. Due to the fixation of a deleterious allele in the entire population of Y chromosomes, after each click there will be a selective advantage for increasing the activity of the corresponding mutation-free X alleles. This could, in principle, lead to the evolution of dosage compensation in a gene-by-gene basis (Charlesworth 1998; Charlesworth and Charlesworth 1997).
1.3.4. Selfing populations

In self fertilizing populations, recombination is effectively restricted. Heller and Maynard Smith (1979) proposed that highly selfing populations would be subject to Muller’s ratchet in the same manner as asexual populations. They calculated the frequency of the individuals free of deleterious mutations ($x_0$) and argued that, if $N_x_0$ is small, the class of homozygous individuals free of mutations may be lost by drift, just as in the haploid case. A click of the ratchet, will therefore occur. They showed that, in the case of complete recessivity, the expected number of individuals in that class will be

$$n_{\text{Self}} = Ne^{-\frac{\mu}{2s}}$$

where the mutation rate per genome per generation is $\mu$, $s$ being the effect of a mutation when homozygous. This expression is identical to the one for the haploid case, because the diploid mutation rate is twice the haploid one. Heller and Maynard Smith concluded that the operation of the ratchet would be the same in a haploid asexual population and in a diploid selfing one. This conclusion is based on the assumption that the size of the least-loaded class is the sole determinant of the speed of the ratchet. In chapter 2, I will show that such a condition is not sufficient for that purpose and that for the same mean number of mutations (and therefore the same $n_0$), large selection coefficients (such as those against the homozygous effects of mutations) can considerably slow down Muller’s ratchet. Simulation studies by Charlesworth et al. (1993b) and Pamilo et al. (1987) of the effects of the ratchet in highly inbred populations seem to point towards the conclusion that the ratchet could have significant effects only in small selfing populations.
### 1.3.5. What stops or slows Muller's ratchet?

Several authors have evaluated the consequences of the violation of the major assumptions classically made in trying to quantify the relevance of the ratchet. From what I have just reviewed two natural questions arise (Bell 1988; Charlesworth et al. 1993b; Maynard Smith 1988): how much recombination and how much outcrossing will stop the ratchet?

Bell (1988) explored the role of recombination in halting the ratchet and proposed an expression to calculate the critical amount of recombination that would protect a population from its operation. The critical mean number of crossovers per chromosome, $r$, proposed is given by:

$$\ln(r) = -1.6 - \ln(n_0) \Leftrightarrow r = \frac{0.2}{n_0}$$

Bell (1988, pg.77). He concluded that in large populations a small amount of recombination would stop the ratchet, but smaller populations would need high rates of recombination in order to do so. However, the simulation work of Charlesworth et al. (1993b) showed that the validity of the above expression is highly dependent on the parameter values. For the parameter space that they analysed, Bell's formula appears to give a good approximation when mutation rates are low, which means that the mean number of mutations per gamete is small. If this is not the case, and if $n_0 << 1$, the critical amount of recombination is much lower than that predicted by Bell's expression. From their simulation results, Charlesworth et al. (1993b) essentially concluded that a very small amount of recombination could stop Muller's ratchet. Similarly from simulations of populations with different selfing rates they concluded
that a small amount of outcrossing is sufficient to avoid the ratchet. To my knowledge, there is currently no reliable theory to predict the value of $r$ or the value of the outcrossing rate for which this occurs.

The conclusions that I have been referring until now have being derived from models that consider that the effect of a new mutation is independent of the presence of other mutations in the same individual. It was shown however (Charlesworth et al. 1993b; Kondrashov 1994) that, if synergistic epistatic interactions between the mutations are common, such that a new mutation will cause a higher decline on the fitness of an individual that already has other mutations, the speed of the ratchet will be lower than in the case of multiplicative fitness. For example it is easy to see that in the extreme case of truncation selection, in which an individual carrying more than a certain number $m$ of mutations does not survive, the ratchet will be stopped after clicking $m$ times. However, Butcher (1995) has also shown that epistasis will not halt the ratchet if a continuous distribution of selection coefficients, with a high density of weak deleterious mutations, is assumed.

The existence of compensatory mutations, that occur in different loci of the genome and compensate for the deleterious mutations already present, can, in the short run, be as effective as recombination in halting the decline in fitness due to the ratchet (Wagner and Gabriel 1990).

Segregation can also strongly decelerate the ratchet. Antezana and Hudson (1997) compared diploid asexual populations to non-recombining populations with segregation. In the first case the ratcheting unit is the whole genome, but in the latter case the ratcheting unit is the non-recombining chromosome. The mutation rate is therefore considerably reduced which will cause a substantial reduction in the speed
of the ratchet. As intuitively obvious, even with no crossing over, the more segmented the genome becomes, the slower is the ratchet.

As it is clear from this brief review, much theoretical work has been done regarding different aspects of this process. But there is still no general analytical solution to the very simple question: what is the rate at which the ratchet operates, or putting it in another way, how much time does it take for the ratchet to click one notch?

Although much effort has been put to find a general expression for the speed of the ratchet, to the present no general solution (which is reasonably accurate over the entire parameter space) has been found, even under the simplest assumptions of multiplicative fitness and equal effects of the deleterious mutations. Two major ways to get approximate expressions for the speed of the ratchet have been pursued. One is to use quantitative genetics theory, and to treat the number of deleterious mutations as a quantitative trait. Calculation of the moments of the distribution allows the prediction of the rate of the ratchet (Higgs and Woodcock 1995; Pamilo et al. 1987). Although the expressions obtained produce reasonable results when considering high mutation rates, they do not give reasonable results when this is not the case. A major difficulty is that it is hard to get a good approximation for the variance of the number of mutations, since this is a function of the third moment of the distribution, which is in turn a function of higher moments, and they do not appear to scale in a simple way with $N$, $U$ and $s$ (Higgs and Woodcock 1995). Another approach is to use diffusion theory to calculate the mean time to loss of the least loaded class, assuming that the mutational classes in population are close to their deterministic equilibrium expectation (Charlesworth and Charlesworth 1997; Stephan et al. 1993). In chapter 2, I choose to use this method and propose an approximation for the time between clicks.
of the ratchet and the decline in mean population fitness. I compare this approximation with previous diffusion approximations, and conclude that it leads to reasonably accurate predictions when compared to the Monte Carlo stochastic simulations.

In chapter 3, I consider a more realistic mutational model. I try to analyse the combined effect of Muller's ratchet and background selection in a haploid asexual population, with a simple model. The model considers the existence of two major types of deleterious mutations, as probably occurs in natural populations. It is shown that the background selection effect, through the quick elimination of strongly deleterious mutations, can considerably accelerate Muller's ratchet, to the point that with these two processes operating, large non-recombining chromosomes may experience loss of gene function in a relatively short amount of evolutionary time, even in large populations.

As with background selection, in chapter 4 it is shown that the operation of the ratchet causes a strong reduction in the level of variability at neutral loci and an approximate equation to calculate the expected level of neutral genetic diversity is presented. But, unlike background selection, the operation of the ratchet can cause a considerable increase in the proportion of low frequency variants in the population, which might be detectable in a sample.

Finally, in chapter 5, I study what the potential effects of a more realistic model for the distribution of selection coefficients may be. In contrast to the models considered in chapter 2, 3 and 4, for the model in chapter 5 I could not get any analytical approximations, but the simulation results suggest that there are circumstances under which the results of chapter 2, 3 and 4 will constitute reasonable approximations.
1.4 Weak Hill-Robertson Effects

Suppose now that many sites in a tightly linked genome are undergoing mutations to very weakly selected alleles, for which \( N_e s \) is around one. Such mutations segregate at much higher frequencies than those involved in background selection and the ratchet, and there is now a non-negligible chance of back-mutation (Comeron et al. 1999; McVean and Charlesworth 2000). An example of this is provided by synonymous mutations affecting codon usage, which may affect the efficiency or accuracy of translation. Models of genomes with many sites subject to such weakly selected mutations have shown that, as linkage increases, both within-species variability and the mean level of adaptation (measured by the frequency with which optimal codons are used) decrease. This effect results from the cumulative effect of numerous polymorphic selected sites on \( N_e \), and is larger, the more sites involved (McVean and Charlesworth 2000). This process may be of importance in the evolution of any sizeable genome where linkage is complete, and in genomic regions with low levels of recombination.

1.5 Relating Models to Observations

From what I have written, several different models make similar predictions. This means that it is difficult to ascribe any given observation to just one process. The question then becomes: are there any conceivable observations that would allow one to identify which process is operating? It turns out that the distribution of frequencies of neutral variants in a sample of sequences- the frequency spectrum- is model-dependent. A large distortion of the frequency spectrum towards rare variants is more likely with selective sweeps (Figure 1.1) than with the other processes (although there are situations in which the operation of Muller’s ratchet can produce similar effects, as will be shown in chapter 4). In addition, when there is some
recombination, so that initially rare variants are not necessarily swept to complete fixation, a transient signature of a sweep is provided by the occurrence of an excess of "derived" variants at a high frequency within a sample (one can infer if a variant is derived or ancestral from the sequence of a closely related species) (Fay and Wu 2000). Other features of genetic variation, such as patterns of linkage disequilibrium, may also be useful in conducting tests of alternative models. Perhaps the best way to quantify the relative importance of these processes is to get solid estimates of the rate at which deleterious mutations occur, and of the distribution of their effects on fitness. Although this is a simple question to ask, it is hard to answer. But given such information, and using the increasing amount of information on DNA sequence variation and evolution, one can perhaps try to answer an even harder question: what is the rate at which advantageous mutations occur, and what are their effects on fitness?
2 The speed of Muller's ratchet

This chapter is based on two articles published in *Genetics* (Gordo and Charlesworth 2000a; Gordo and Charlesworth 2000b)

2.1 Introduction

New mutations arise continuously within populations, the vast majority probably being slightly deleterious (Crow 1993). The presence or absence of recombination in a population has a very strong impact on the dynamics of the deleterious mutations which it contains. With free recombination, deleterious mutations can be eliminated effectively but, as discussed in the previous chapter, in the absence of recombination, the efficacy of selection is reduced and deleterious mutations can accumulate in finite populations (Felsenstein 1974; Fisher 1930, p. 122; Hill and Robertson 1966; Muller 1964).

Consider first an effectively infinite population subject to recurrent deleterious mutations. Such population will achieve an equilibrium resulting from the continuous appearance of new mutations opposed by selection against them: deterministic mutation-selection balance. With independent and identical effects of each mutation, the equilibrium number of mutations in a haploid randomly mating population follows a Poisson distribution with mean $U/s$ (Haigh 1978; Kimura and Maruyama 1966), where $U$ is the per genome mutation rate and $s$ is the selection coefficient against a deleterious mutation. At this deterministic equilibrium, the number of individuals free of mutations (the best or “least-loaded” class) in a large population of $N$ breeding adults is $n_0 = N \exp(-U/s)$. This is true whether the population is sexual or asexual, as long as it is sufficiently large.

But in any finite population, random genetic drift plays a role and may perturb this equilibrium, leading to the loss of the best class. In the absence of
recombination, and with the reasonable assumption that back-mutation is negligible for strongly selected mutations (those with \( Ns > 1 \)), the loss of this class is irreversible and mutations will continually accumulate in the population, leading to the decline of its mean fitness. This is the process known as Muller's ratchet (Muller 1964). Once the best class is lost, the new least-loaded class is now the one that has one mutation, but this is also subject to stochastic loss, so that a repetition of successive losses of the least-loaded classes can be seen as successive clicks of the ratchet.

One of the important questions in this process concerns the rate or speed at which it operates, or: how much time does it take for the ratchet to click one notch? As I have discussed, the degeneration of the \( Y \) chromosome (Charlesworth 1978; Charlesworth 1996b; Charlesworth and Charlesworth 1997; Charlesworth and Charlesworth 1998; Rice 1994) and the fate of asexual populations (Gabriel et al. 1993; Lynch et al. 1993; Lynch et al. 1995; Lynch and Gabriel 1990; Pamilo et al. 1987) may involve the operation of Muller's ratchet, so that the quantification of its rate is of great biological importance. This chapter deals with such a calculation. Here I study the simplest possible mathematical model, and put particular emphasis on parameter values that are possibly relevant to the evolution of \( Y \) chromosomes.

Haigh (1978) suggested that the most important parameter for the ratchet mechanism would be \( n_o \), because it is the loss of the best class that drives the process; the smaller \( n_o \), the faster the ratchet is likely to be. Although Haigh suggested an expression (equation (9a) in Haigh (1978)) for the time between clicks of the ratchet, this is basically a fit to his simulation results. Bell (1988) also suggested an expression based on a fit of the time to extinction of the best class as a function of \( n_o \) (roughly \( 10n_o \)). Several attempts towards the quantification of the process have subsequently been made, either using a quantitative genetics approach for estimating the rate of change of the average number of mutations (Gabriel et al.
1993; Higgs and Woodcock 1995; Lynch et al. 1993; Pamilo et al. 1987; Prugel-Bennett 1997), or using diffusion theory to calculate the mean time to loss of the least-loaded class (Charlesworth and Charlesworth 1997; Stephan et al. 1993).

Below it is shown that the size of the best class is not sufficient to predict the speed of the ratchet. For the same value of \( n_0 \), the ratchet can turn at very different speeds. In section 2.2 an approximation based on a diffusion equation, for the case \( n_0 > 1 \), which may apply to the evolution of the \( Y \) chromosome (Charlesworth 1996b), is presented. In section 2.4, Monte-Carlo simulations of asexual haploid populations are done to check the validity of the diffusion approximation, under the parameter range for which \( n_0 > 1 \). For the cases when \( n_0 < 1 \), Gessler (1995) has derived an approximation that seems to work well. Comparisons between the simulation results and the predictions from the analytical expression suggest that the formulae seem to predict the rate of the process better than the previous formulae, for a region of parameter space which is of biological interest (Charlesworth and Charlesworth 1997), namely for values of selection coefficients of a few percent as estimated from classical mutation accumulation experiments (Crow 1993; Keightley 1994; Keightley 1996; Keightley and Eyre-Walker 1999; Keightley and Ohnishi 1998). In section 2.5, I refine this approximation for the cases when \( s \) is small, because recently it has been suggested that there may be a large class of small effect mutations that is not detectable in the classical mutation accumulation experiments (Davies et al. 1999).

### 2.2 Approximation based on the diffusion equation

Consider a haploid asexual population at equilibrium under mutation-selection balance, with \( x_0 = \exp(-U/s) \) being the frequency of individuals in the least-loaded class. The existence of this equilibrium requires that the number of individuals in the
best class \( (n_0 = N x_0) \) be greater than one. When \( n_0 < 1 \), this equilibrium may not be approached in a finite population, because it requires the existence of individuals that have a very low probability of actually being present (Gessler 1995; Gessler and Xu 1999).

The way the frequency of the best class varies through time is dictated by mutation taking it below the equilibrium value, selection restoring it to that value, and by stochastic fluctuations due to drift. One wants to quantify how much time it takes for the frequency of the best class (from now on denoted by \( x \)), with initial value \( x_0 \), to reach the value zero. One way to do this is to use a diffusion equation for the density function of the time until absorption occurs, subject to the condition that \( x = 0 \) is the only absorbing state (Ewens 1979, equations 4.39, 4.40, p. 123). In order to solve this equation, one has to evaluate the deterministic change (drift coefficient) and stochastic variance (diffusion coefficient) in \( x \).

Assuming a Wright-Fisher population, the diffusion coefficient is just the variance due to binomial sampling of \( N \) individuals from the previous generation (Charlesworth and Charlesworth 1997; Stephan et al. 1993):

\[
b(x) = \frac{x (1-x)}{N} \approx \frac{x}{N} \tag{1a}
\]

assuming \( x << 1 \).

The drift coefficient, representing the expected change in \( x \) due to mutation and selection (Charlesworth and Charlesworth 1997; Stephan et al. 1993), is:

\[
a(x) = \frac{x (e^{-w} - \bar{w})}{\bar{w}} = x \frac{\Delta \bar{w}}{\bar{w}} \tag{1b}
\]

where \( \bar{w} \) is the mean fitness of the population and \( \Delta \bar{w} \) is the difference between the current mean fitness and the mean fitness at equilibrium.

Let us now make the simplifying assumption that the changes in mean fitness are sufficiently small that they can be approximated by small perturbations from the
equilibrium value (Stephan et al. 1993). This implies that the system is close to its equilibrium state most of the time, as is supported by our simulations (see figure 2.9). I will assume that the perturbations in $\tilde{w}$ are mostly due to small fluctuations in the least-loaded class. This assumption was employed by Stephan et al. (1993) and Charlesworth and Charlesworth (1997), and is justified in practice by the observation that, for large $N$, the distribution among the classes that are present at any time remains close to the Poisson distribution given by the deterministic equilibrium formula (see Table 1 and figure 4 of Charlesworth and Charlesworth 1997). One may express the mean fitness close to equilibrium as a Taylor expansion in $x/x_0$:

$$
\tilde{w}\left(\frac{x}{x_0}\right) = \tilde{w}_{eq} + \left[ \frac{\partial \tilde{w}}{\partial (x/x_0)} \right]_{eq} \left( \frac{x}{x_0} - 1 \right) + \mathcal{O}\left( \frac{x}{x_0} - 1 \right)^2 \quad (2a)
$$

Taking just the linear term in (2a), one obtains an approximation for the reduction in mean fitness below its equilibrium value as:

$$
\Delta \tilde{w} = K \left( 1 - \frac{x}{x_0} \right) \quad (2b)
$$

as previously assumed by Stephan et al. (1993), where

$$
K = x_0 \left[ \frac{\partial \tilde{w}}{\partial x} \right]_{eq}
$$

When the frequency $x$ is above its equilibrium value, the system responds with a reduction of mean fitness ($\Delta \tilde{w} < 0$) towards the equilibrium value, and when the opposite happens, so that $x$ goes below $x_0$ ($\Delta \tilde{w} > 0$), the system responds with an increase in $\tilde{w}$. The forces underlying the response of the system towards equilibrium are selection (when $x < x_0$), and mutation (when $x > x_0$), which are parameterized by
$K$ in our small-perturbation model. $K$ can be estimated as follows. If by chance the least-loaded class goes extinct ($x = 0$), then the ratchet has clicked and the mean fitness will decline towards a new deterministic equilibrium value, $(1 - s) e^{-U}$. Then the net loss of mean fitness due to a click of the ratchet would be given by $\Delta w = s e^{-U}$ if the distribution instantaneously recovered its Poisson equilibrium with a new least-loaded class with abundance $n_0$, after the loss of the least-loaded class. In practice, stochastic fluctuations mean that this equilibrium is never achieved exactly.

In fact, the approach to the neighbourhood of equilibrium takes some time (Haigh 1978), and just after a click the new best class is above its equilibrium value, so that $\Delta w < s e^{-U}$. Haigh (1978) showed that, after a click, the new least-loaded class rapidly approaches a value close to $1.6 n_0$, and then the approach to the new equilibrium value is slower (Haigh 1978, pg 255). If this is the case, equation (2b) means that the reduction in mean fitness after a click will not be $s e^{-U}$ but approximately $0.6 s e^{-U}$. I thus set $K = 0.6 s e^{-U}$ and we test how accurate this approximation is with the help of simulations (see simulation results below).

The drift coefficient may now be written as:

$$a(x) = 0.6 s \left( 1 - x/x_0 \right) x$$ (2c)

Using these drift and diffusion coefficients, the time spent in the frequency interval $[0, x_0]$ (equation 4.39, Ewens 1979) is:

$$T_{0,x_0} = \int_0^{x_0} \frac{2N}{x} G(x) \left\{ \int_0^{x} G(x') \, dx' \right\} \, dx$$ (3a)

and the time spent in the interval $[x_0,1]$ (equation 4.40, Ewens 1979) is:

$$T_{x_0,1} = \int_{x_0}^{1} \frac{2N}{x} G(x) \left\{ \int_0^{x} G(x') \, dx' \right\} \, dx$$ (3b)
where

\[ G(\xi) = \exp \left[ -2 \int_{0}^{\xi} \frac{a(z)}{b(z)} \, dz \right] = \exp \left[ \frac{2N}{x_0} 0.6s \left( \frac{\xi}{2} - x_0 \right) \right]. \]

Using expressions (3a) and (3b), and evaluating the integrals numerically for a given population size, mutation rate and selection coefficient, the expected time to loss of the least loaded class is obtained as:

\[ T(N,U,s) = T_{x_0} + T_{x_0,1}. \]

This expression differs from Stephan et al. (1993) and Charlesworth and Charlesworth (1997) diffusion approximations in the following aspects: first the approximation of the drift coefficient leading to equation (14) of Stephan et al. is slightly different from equation (2c), although it becomes very similar to equation (2c) with \( K = 0.5s e^{-U} \). Second, and more importantly, in Stephan et al. equation (14) the time spent above \( x_0 \), corresponding to \( T_{x_0,1} \) in the approximation just presented, is zero, since they treat \( x_0 \) as the reflecting barrier. In Charlesworth and Charlesworth diffusion approximation the time spent in \( T_{x_0,1} \) is considered, but the approximation for the drift coefficient is different from the one just presented (compare their equation (A4a) with equation (2c)).

2.3 Simulation methods

For a given population size \( (N) \), genomic mutation rate to deleterious mutations \( (U) \), and selection coefficient against each mutation \( (s) \), haploid asexual populations were simulated starting at mutation-selection equilibrium (Kimura and Maruyama 1966), i.e. the number of individuals in the class with \( m \) mutations is
Assuming that the sequence of events is mutation, reproduction and selection, populations were then run for 100 generations. After this initial period, populations were run for more than 2,000 generations, and up to 100,000 generations for conditions under which the ratchet clicks slowly, so that the average time between clicks of the ratchet could be measured. Every generation, the number of mutations in every individual is counted and the number of individuals with the least number of mutations (least-loaded class) is recorded. If, at a given generation, the number of mutations in the least-loaded class increases, the ratchet has clicked. To form a new generation, individuals are sampled randomly from the previous generation, then subjected to the occurrence of mutations sampled from a Poisson distribution with mean $U$, and assigned probabilities of survival as $(1-s)^k$, where $k$ is the number of mutations that an individual carries. A new generation of $N$ individuals is constructed by comparing the probability of survival of each individual with a pseudo-random number drawn from a uniform distribution in the interval $[0,1]$. Each run was repeated several times; generally 5 replicates were performed in order to obtain the results presented in the next section.

Although this simulation procedure does not follow the fate of each mutation at a particular locus, which is extremely time consuming, it gives the same results as the multi-locus stochastic simulations of Charlesworth and Charlesworth (1997) as far as the estimation of the time between clicks of the ratchet is concerned, for all parameter sets tested (results not shown).
2.4 Simulation results

If the loss of the least loaded class is the determining factor in driving the ratchet (Haigh 1978; Bell 1988; Maynard Smith 1978), one would expect that the time between clicks of the ratchet would stay approximately constant over a range of parameter values that keep \( n_0 \) constant. Figures 2.1, 2.2 and 2.3 show the simulation results for several parameter sets chosen such that \( n_0 \) stays constant. In figure 2.1, \( s \) is kept constant at 0.015, and \( N \) changes with \( U \) to keep \( n_0 \) constant (either 20.2 or 202).

**Figure 2.1** Average time between clicks of the ratchet, \( \pm 2 \) standard errors (SE), for constant \( s \) as a function of the population size (as \( N \) increases, \( U \) increases so that \( n_0 \) is maintained constant; \( s \) is set equal to 0.015). \( n_0 \) is 202 for circles and 20.2 for triangles. The points joined with a line are the estimated times given by the approximation presented in the text.

Figure 2.1 shows that the time between clicks of the ratchet does not change significantly, over an order of magnitude change in \( N \). The increase in \( N \) seems to be compensated by the increase in \( U \).
**Figure 2.2** Average time between clicks of the ratchet, ± 2 SE, for constant $U$ as a function of the population size (as $N$ increases, $s$ decreases so that $n_0$ is maintained constant; $U$ is set equal to 0.1). $n_0$ is 202 for circles and 20.2 for triangles.

**Figure 2.3** Average time between clicks of the ratchet, ± 2 SE, for constant $n_0$ as a function of the selection coefficient (as $s$ increases, $U$ increases so that $U/s$ is constant). $N=30,000$ for circles ($n_0=202$) and 3,000 for triangles ($n_0=20.2$).
In figure 2.2, $U$ is kept constant, at 0.1, and $N$ changes together with $s$ to keep $n_0$ constant (with the same values as in figure 2.1). Although for small values of $n_0$ there seems to be no significant difference in the speed of the ratchet over an order of magnitude change in population size, for a higher value of $n_0$ the difference is evident: as $s$ becomes large (small $N$ in the plot), the speed of the ratchet is greatly reduced. For example, for a population size of 705 individuals, with a selection coefficient of 0.08, I did not observe any click over 50,000 generations, but for a population of 19,000 individuals, with $s = 0.022$, the average time for a click is about 1,560 generations.

In figure 2.3, $N$ is kept constant and $U$ changes with $s$ (the mean equilibrium number of mutations, $U/s$, has the value 5), so that $n_0$ is constant. One sees that, for either small $U$ and small $s$ or for large $U$ and large $s$, the speed of the ratchet is greatly reduced. These results show that, as noted previously by Stephan et al. (1993), the size of the best class is not sufficient to predict the speed of the ratchet, because the ratchet can turn at very different speeds for a constant $n_0$. One observes that for the same $n_0$, keeping $N$ constant and varying $U$ and $s$ so that $U/s$ is constant, there is a value of $U$ and $s$ for which the time of the ratchet has a minimum, i.e., for the same $n_0$, increasing $s$ can both slow down and speed up the ratchet (see the U-shaped curves in figure 2.3). This shape is not predicted by any of the previous formulae. In contrast, such a minimum is predicted by the approximation presented here, although in the region of very small $U$ and very small $s$ the approximation gives lower times than the simulations. It also underestimates the time between clicks of the ratchet for small population size (or small $n_0$).

The possible reasons why this minimum is observed, and why the present approximation underestimates the time for small population size, deserve some comment. Other things being equal, decreasing $s$ should speed up the ratchet and decreasing $U$ should slow it down. In the case of figure 2.3, both $s$ and $U$ are
changing in order to keep $U/s$ constant, so that a minimum may occur, due to the fact that the dependence of the time on the mutation rate is different from that on the selection coefficient. In the region where $s$ is very small, so that each mutation has very little effect on fitness, $U$ is also very small. This means that the probability of a mutation occurring is very small, and the force of mutation that drives individuals from the best class to the next class is greatly reduced, leading to a slower ratchet. In the region where the mutation rate is large, the selection coefficient is also large, so that although mutations keep appearing at a high rate, selection is so efficient in restoring the best class that a great number of individuals come from the least-loaded class, which leads to a slower ratchet.

One observes from the comparison of the theoretical formula and the simulation results that, as long as $s$ is not extremely small or large (figures 2.1 - 2.3), the predictions seem to approximate the simulations reasonably well, especially when $N$ is big (or $n_0$ is large). If $N$ is small, it is more difficult for the system to maintain itself close to equilibrium, because drift is dominating. In this case, an approximation based on small perturbations may become inadequate. If $s$ is small, the diffusion approximation consistently underestimates the average time obtained in the simulations. I will try to obtain a better approximation for these cases in section 2.5.

I now ask how the speed of the ratchet changes with population size, for a given mutation rate and a constant selection coefficient. The simulation results and the expected times calculated with the various approximations discussed above are shown in figures 2.4, 2.5 and 2.6, for different values of $N$, $U$, and $s$, as a function of $n_0$. The mutation rate and selection coefficient were chosen to lie in the parameter range which may be most relevant to the problem of the evolutionary degeneration of an incipient $Y$ chromosome or neo-$Y$ chromosome (Charlesworth 1978; Charlesworth 1996b; Charlesworth and Charlesworth 1997). The values of $U = 0.04$
and $s = 0.01$ might be considered reasonable for the case of a Drosophila chromosome arm, if widely accepted estimates of mutational parameters are used (Charlesworth and Charlesworth 1998; Drake et al. 1998).

Figure 2.4 Mean time between clicks of the ratchet (Trat), ± 2 standard errors, as a function of $n_0$ (this is varied by varying $N$). The mutation rate is 0.075 and the selection coefficient is 0.02. The results of the approximations of Charlesworth and Charlesworth are labelled Tc (Charlesworth and Charlesworth 1997) and of Stephan et al. are labelled Ts and Ts14 for their equations (8) and (14), respectively (Stephan et al. 1993). The result from the approximation presented here is labelled Tp.
Figure 2.5 The same as in figure 2.4, but for a mutation rate of 0.05 and selection coefficient of 0.015.

Figure 2.6 The same as in figure 2.4, but for a mutation rate of 0.04 and a selection coefficient of 0.01.
In each figure, the time between clicks of the ratchet is plotted against \( n_0 \), for a fixed value of \( U \) and \( s \), so that an increase in \( n_0 \) is solely due to an increase in \( N \). The population size varies in the interval \([1,000; 13,500]\) in figure 2.4, \([1,000; 10,000]\) in figure 2.5 and \([1,000; 30,000]\) in figure 2.6. One observes that the approximation of Charlesworth and Charlesworth (Tc), contrary to what was previously thought (Charlesworth and Charlesworth 1997; Orr and Kim 1998), greatly underestimates the speed of the ratchet for large population sizes in the parameter range considered here. Although Stephan et al. could not establish exactly the range of validity of their two approximations, they suggested use of their equation (8) for predicting the speed of the ratchet for the range of selection coefficients considered here. From comparison with the simulations presented, one can see that their equation (14) seems to describe the rate of change with \( N \) of the time between clicks of the ratchet better than their equation (8), although it always underestimates the absolute time in this parameter range.

In simulations done to check the change in the ratchet’s speed with different mutation rates (figure 2.7), one can see that the range of parameters for which Stephan et al. equation (14) gives a better quantification of the process than their equation (8) is not only dependent on a large population size and strong selection, but is also modulated by the mutation rate. In simulations done to check the change of the ratchet’s speed with different \( s \) (figure 2.8) one sees that, for a given \( N \) and \( U \), as \( s \) increases, the estimated time given by their equation (14) becomes an underestimate. In fact, Stephan et al. studied the process by dividing it in two separate phases: the establishment phase, during which the new least-loaded class reaches a value close to that of the deterministic equilibrium, and the extinction phase, during which the least-loaded class becomes extinct. Their equation (14) is based on the assumption that the population size is large enough that the change in mean fitness of the population is sufficiently small to be approximated by small
Figure 2.7 The time between clicks of the ratchet, ± 2 SE as a function of the mutation rate \((N = 10,000 \ s = 0.02)\).

Figure 2.8 The time between clicks of the ratchet, ± 2 SE, as a function of the selection coefficient \((N = 10,000 \ U = 0.075)\).
perturbations and that selection is strong enough that the process is mostly trapped in
the extinction phase. This assumption is not supported by the Monte-Carlo
simulations (see figure 2.9). The time spent in the establishment phase is the one
above the dashed line; the time spent in the extinction phase is the one below this
line.

Figure 2.9 The fluctuations in the size of the least-loaded class through time, for \(N =
10,000\), \(U = 0.075\), and \(s = 0.02\), for a single simulation run. The dashed line
indicates the equilibrium value of the least-loaded class.

When deriving the approximation, I assumed that the net loss of mean fitness
due to a click of the ratchet would be \(K = 0.6 s e^{-U}\), because the system did not
recover its new equilibrium instantaneously. In figure 2.10A, I plot the mean fitness
of the population as a function of \(1 - x/x_0\) for a set of simulation runs with parameters
\(N = 10,000\) \(U = 0.04\) \(s = 0.01\), after a click of the ratchet. As I assumed that the
mean fitness of the population could be approximated as a linear function of \(1 - x/x_0\),
the slope of the linear regression line plotted in the figure corresponds to the value of
\(K\) in the theoretical approximation. The value assumed in the derivation \(K =\)
5.76x10^3) agrees well with the one from the regression. However, the agreement is not good for a population size of only 1,000 (figure 2.10B).

![Graph A](image1)

\[ y = -0.0053x + 0.951 \]
\[ R^2 = 0.7107 \]

![Graph B](image2)

\[ y = -0.0012x + 0.9551 \]
\[ R^2 = 0.4811 \]

**Figure 2.10** Population mean fitness as a function of 1-x/x₀ for a set of simulation runs with parameters, \( U = 0.04, s = 0.01, N = 10,000 \) (in A) and \( N = 1,000 \) (in B). The line plotted is a linear regression line whose slope represents the value of \( K \) assumed in the theoretical approximation.
This can be attributed at least to two factors: either the population size is so small that one cannot approximate the changes in mean fitness by small perturbations and/or the value of $K$ is different from the one assumed. If I calculate the time by substituting the value of $K$ from the linear fit in the theoretical expression, I find that the time obtained is still below the one measured in the simulations, so that an incorrect value for $K$ is not the only source of the discrepancy.

2.5 The time between clicks and the decline in mean fitness

Although the diffusion approximation just presented for quantifying the mean time between turns of the ratchet seems to work reasonably well for values of $s$ consistent with the classical estimates from mutation accumulation experiments (of the order 1-2%) (Crow 1993) it underestimates the time for much smaller values of $s$ (see figure 2.8). Since the values of $s$ obtained from mutation accumulation experiments are generally overestimates, and it has recently been proposed that there may be a large class of mildly recessive deleterious mutations with selection coefficients much less than 1-2% (Keightley and Eyre-Walker 1999), one needs to have an approximation to the speed of the process which also works in this parameter range.

In this section the previous approximation is re-examine, and I suggest a more robust prediction that seems to work better over a wider range of parameters. In section 2.2 I have tried to determine the speed of Muller's ratchet by modelling the ratchet as a one-dimensional diffusion process, for which the mean time to absorption of the frequency of the least-loaded class was calculated. When deriving the drift coefficient, I assumed small perturbations around the state given by the deterministic mutation-selection balance and also assumed that, just after a turn of the ratchet, the size of the new least-loaded class would rapidly approach a value
close to $1.6 n_0$ (as suggested by the corollary of Haigh's Theorem 1, 1978). When $s$ is intermediate, this is likely to be true, but the smaller the value of $s$ the longer will be the time to approach this value. In particular, when $n_0$ and $s$ are small, then this relaxation time becomes the main determinant of the total time to absorption. It is this additional time ($T_a$), which was assumed previously to be effectively zero, that must be added in order to get a better prediction.

$T_a$ can be approximated by the time it takes to get from the size of the new least-loaded class immediately after one click (which at this point has a value of around $n_1 = n_0 U/s$) to $1.6n_0$ using Haigh's theorem 1 (or equation (3) in Stephan et al. (1993)). Haigh's theorem 1 describes the behaviour of the new least-loaded class after a click as:

$$n_0(t) = n_1 \frac{(1 - s)^t}{1 - e^{-U(1-s)t}}$$

For small $s$, this will be approximately

$$n_0(t) = n_1 (1 - st)$$

and the time ($T_a$) for the least loaded class to get from $n_1=n_0 U/s$ to $1.6n_0$ will be approximately:

$$T_a = \frac{1}{s} \left( \frac{1 - 1.6s}{U} \right)$$

(4)

Therefore the mean time for a turn of the ratchet is:

$$T(N,u,s) = T_a + T_{0,x_0} + T_{x_0,1}$$

(5)

where $T_{0,x_0}$ is the time spent in the frequency interval $[0,x_0]$ and $T_{x_0,1}$ is the time spent in the interval $[x_0,1]$, given by equations (3a) and (3b) in section 2.2.
Table 2.1

Comparison of the mean time between turns of the ratchet between simulations $T$ (with 2 standard errors), and the analytical expression $T(N,U,s)$

<table>
<thead>
<tr>
<th>$N$</th>
<th>$U$</th>
<th>$s$</th>
<th>$n_0$</th>
<th>$T$ (2SE)</th>
<th>$\bar{T}/\bar{w}_i$</th>
<th>No. of fixations$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>275</td>
<td>0.06</td>
<td>0.15</td>
<td>5</td>
<td>63 (5)</td>
<td>66</td>
<td>8.06 x10^-53</td>
</tr>
<tr>
<td>565</td>
<td>0.1</td>
<td>0.03</td>
<td>20</td>
<td>119 (2)</td>
<td>101</td>
<td>3.29 x10^-55</td>
</tr>
<tr>
<td>1000</td>
<td>0.04</td>
<td>0.01</td>
<td>18</td>
<td>169 (39)</td>
<td>144</td>
<td>1.22 x10^-13</td>
</tr>
<tr>
<td>3000</td>
<td>0.0075</td>
<td>0.0015</td>
<td>20</td>
<td>489 (42)</td>
<td>571</td>
<td>2.15 x10^-1</td>
</tr>
<tr>
<td>3000</td>
<td>0.1</td>
<td>0.02</td>
<td>20</td>
<td>135 (14)</td>
<td>120</td>
<td>3.19 x10^-3</td>
</tr>
<tr>
<td>5000</td>
<td>0.04</td>
<td>0.01</td>
<td>92</td>
<td>514 (99)</td>
<td>442</td>
<td>5.68 x10^-5</td>
</tr>
<tr>
<td>6065</td>
<td>0.03</td>
<td>0.005</td>
<td>15</td>
<td>215 (25)</td>
<td>225</td>
<td>8.66 x10^-6</td>
</tr>
<tr>
<td>10000</td>
<td>0.075</td>
<td>0.01</td>
<td>5.5</td>
<td>126 (12)</td>
<td>109</td>
<td>4.78 x10^-18</td>
</tr>
<tr>
<td>13500</td>
<td>0.075</td>
<td>0.02</td>
<td>317</td>
<td>4574 (1913)</td>
<td>4712</td>
<td>1.10 x10^-1</td>
</tr>
<tr>
<td>30000</td>
<td>0.005</td>
<td>0.001</td>
<td>202</td>
<td>1743 (225)</td>
<td>1601</td>
<td>6.98 x10^-1</td>
</tr>
<tr>
<td>30000</td>
<td>0.01</td>
<td>0.002</td>
<td>202</td>
<td>1225 (272)</td>
<td>1200</td>
<td>4.42 x10^-1</td>
</tr>
<tr>
<td>30000</td>
<td>0.1</td>
<td>0.02</td>
<td>202</td>
<td>1543 (275)</td>
<td>1521</td>
<td>1.44 x10^-3</td>
</tr>
<tr>
<td>30000</td>
<td>0.175</td>
<td>0.035</td>
<td>202</td>
<td>4890 (454)</td>
<td>3860</td>
<td>2.62 x10^-2</td>
</tr>
<tr>
<td>50000</td>
<td>0.04</td>
<td>0.01</td>
<td>918</td>
<td>45922</td>
<td>38957</td>
<td>8.96 x10^-1</td>
</tr>
<tr>
<td>500000</td>
<td>0.015</td>
<td>0.0025</td>
<td>1239</td>
<td>4924 (764)</td>
<td>7785</td>
<td>7.76 x10^-1</td>
</tr>
<tr>
<td>500000</td>
<td>0.015</td>
<td>0.0015</td>
<td>23</td>
<td>732 (66)</td>
<td>1343</td>
<td>3.59 x10^-1</td>
</tr>
<tr>
<td>500000</td>
<td>0.03</td>
<td>0.003</td>
<td>23</td>
<td>490 (114)</td>
<td>716</td>
<td>4.66 x10^-2</td>
</tr>
<tr>
<td>500000</td>
<td>0.3</td>
<td>0.03</td>
<td>23</td>
<td>167 (79)</td>
<td>139</td>
<td>4.29 x10^-38</td>
</tr>
<tr>
<td>500000</td>
<td>0.03</td>
<td>0.005</td>
<td>1239</td>
<td>12819 (5681)</td>
<td>17625</td>
<td>8.22 x10^-1</td>
</tr>
<tr>
<td>500000</td>
<td>0.04</td>
<td>0.004</td>
<td>23</td>
<td>371 (86)</td>
<td>570</td>
<td>4.51 x10^-3</td>
</tr>
<tr>
<td>500000</td>
<td>0.04</td>
<td>0.005</td>
<td>168</td>
<td>940 (106)</td>
<td>1018</td>
<td>6.95 x10^-2</td>
</tr>
<tr>
<td>500000</td>
<td>0.04</td>
<td>0.006</td>
<td>636</td>
<td>3868 (322)</td>
<td>4657</td>
<td>4.59 x10^-1</td>
</tr>
<tr>
<td>500000</td>
<td>0.04</td>
<td>0.0065</td>
<td>1063</td>
<td>13381 (6210)</td>
<td>19225</td>
<td>7.84 x10^-1</td>
</tr>
<tr>
<td>500000</td>
<td>0.05</td>
<td>0.0075</td>
<td>636</td>
<td>3912 (369)</td>
<td>5910</td>
<td>3.82 x10^-1</td>
</tr>
<tr>
<td>500000</td>
<td>0.056</td>
<td>0.008</td>
<td>456</td>
<td>2420 (303)</td>
<td>3219</td>
<td>1.90 x10^-1</td>
</tr>
<tr>
<td>500000</td>
<td>0.065</td>
<td>0.01</td>
<td>752</td>
<td>9310 (1306)</td>
<td>16934</td>
<td>5.83 x10^-1</td>
</tr>
<tr>
<td>500000</td>
<td>0.07</td>
<td>0.01</td>
<td>456</td>
<td>2711 (411)</td>
<td>4022</td>
<td>1.57 x10^-1</td>
</tr>
<tr>
<td>500000</td>
<td>0.08</td>
<td>0.01</td>
<td>168</td>
<td>812 (113)</td>
<td>915</td>
<td>2.05 x10^-3</td>
</tr>
</tbody>
</table>

$^a$ The expected ratio of the mean fitness of the population and the initial mean fitness after 500,000 generations. $^b$ Expected number of fixations of deleterious mutations after 500,000 generations.
While $T_a$ constitutes the deterministic time for the frequency to approach a state close to the new mutation-selection balance, the other terms constitute the mean time of the stochastic process leading to absorption. For a given $N$ and $U$, small values of $n_0$ correspond to small values of $s$, and $T_a$ dominates the other terms, but as $s$ increases so does $n_0$, and the value of $T_a$ becomes less relevant compared with the other terms. In Table 2.1 I compare the results of this formulation with the ones obtained by simulations. The simulation method is as described in section 2.3. The parameter values were chosen as follows. The large values of $N$ in Table 2.1 are of considerable biological importance, since we want to analyse the role of Muller's ratchet in degeneration of Y and neo-Y chromosomes in systems such as *Drosophila miranda*, for which the effective population size is thought to be in the order of hundreds of thousands or millions (Yi 2000; Yi and Charlesworth 2000). The values of the deleterious mutation rate were chosen to cover a region that is possibly reasonable in the light of various mutation accumulation experiments (Keightley and Eyre-Walker 1999). All the parameters are constrained to the condition $n_0 > 1$ (see Gessler 1995 for $n_0 < 1$). For all the simulation results reported in the previous section and for other parameters tested, the new expression performs better than the old one. In particular, the discrepancy in the time between clicks observed in the U-shaped curve of figure 2.3, for small $s$, is accounted for with the new expression.

In figure 2.11 I show, as an example, the change of the size of the least loaded-class over time intervals of 10 generations after a turn of the ratchet, taken from several simulation runs. The parameter values are: $N = 10,000$, $U = 0.03$ and $s = 0.005$. With these parameters, $n_0 = 25$, $1.6$ $n_0 = 40$ and $n_1 = 149$. Although there is an enormous variance in the behaviour of the change in size of the least-loaded class, on average (bold line in the figure) the behaviour is close to what has been assumed. Immediately after a turn of the ratchet the size of the least-loaded class is close to $n_1$,
then in 100 to 200 generations it approaches a value close to \( n_0 \). This pattern is essentially the same for other parameter values that were tested.

![Diagram showing the change in size of least-loaded class after a turn of the ratchet](image)

**Figure 2.11.** Change in size of least-loaded class after a turn of the ratchet has occurred (at generation 0) for different simulation runs (correspondent to different symbols). The bold line is the average size.

One fact is probably worth noting: although one can, with one expression, estimate reasonably well the time between turns of the ratchet - over very different values of \( N, U \) and \( s \) - from a few generations to thousands of generations, when \( s \) is large (>0.04) the expression underestimates the time obtained in the simulations. In this range, none of the diffusion approximations is accurate, as is expected from the conditions for diffusion theory to be reliable (Ewens 1979, chap. 4).

Under this model, for which each mutation causes an identical and independent deleterious effect on fitness, the decline in the logarithm of mean fitness is:

\[
\frac{\partial \ln w}{\partial t} = \frac{\ln(1-s)}{T} \approx -\frac{s}{T} \quad \text{for small } s
\]  

(5)
where $T$ is the mean time for a click.

Clearly, stronger deleterious mutations cause a bigger decline in log mean fitness per click but take more time to accumulate, while weaker deleterious mutations will accumulate faster but cause a smaller decline in log mean fitness, as noted before by Lynch et al. (1993). Therefore there is a value of $s$, say $s_{\text{max}}$, that maximises the decline in log mean fitness. The partial derivative of $s/T$ with respect to $s$ will be zero at this point. I can easily calculate, using the approximation just presented, the approximate expected value of $s_{\text{max}}$. This value is obviously a function of $N$ and $U$. For example, in the case of a large population with $N = 500,000$, if the mutation rate is 0.04, then mutations whose effect is around 0.004 are expected to cause the biggest decline in log mean fitness ($s/T \sim 1.3 \times 10^{-5}$). If $U$ is smaller, say 0.02, then weaker mutations will correspond to the maximum rate of decline, but cause a much lower rate of decline ($s/T \sim 3.7 \times 10^{-6}$) than in the first case.

From equation (5), the ratio between the mean fitness at any time and the initial mean fitness of the population before any turn of the ratchet can be calculated. I show this ratio after 500,000 generations in Table 2.1. I also show the expected number of fixed deleterious mutations after this time, since it is known that, in the long run, the rate of the ratchet is the rate of fixation of deleterious mutations (Charlesworth and Charlesworth 1997; Higgs and Woodcock 1995).

2.6 Discussion

The equilibrium size of the least loaded class, $n_0$, is usually regarded as the chief parameter that determines the speed of the ratchet (Bell 1988; Charlesworth and Charlesworth 1998; Gessler 1995; Haigh 1978). The simulations presented here show that this parameter is not sufficient for this purpose, as noted previously by
Stephan et al. (1993). For example, for an equilibrium size of the best class of around 200 we can get average times for one click of the ratchet varying from 900 to 8000 as a result of changes in selection coefficient (0.015-0.04) and mutation rate (0.075-0.2).

The parameter space considered here has been constrained to ensure $n_o > 1$, so that the Poisson distribution expected under mutation-selection balance can be approximately attained. Gessler (1995) has shown that this balance will not be met for conditions under which $n_o < 1$, because selection is too weak to counter mutation pressure. In this case, the distribution of the number of mutations is not Poisson but is close to a shifted negative binomial distribution, whose parametrization allows an estimation of the rate of the ratchet. I will turn to this parameter range in the next chapter.

For $n_o \gg 1$, the approximation presented here for the advance of the ratchet based on a diffusion equation seems to give a better prediction of the time between clicks of the ratchet than the previous approximations, for moderate selection coefficients in a range that is compatible with the biological data, provided the population size is not too small, so that the stochastic fluctuations are not too violent. As noted before by Stephan et al. (1993), for intermediate selection coefficients the establishment phase and the extinction phase are blurred and a separate analysis of these two phases does not well predict the outcome. An approximation based on the assumption that the mean fitness of the population is affected solely by fluctuations of the least loaded class ($\tilde{w} = (x - x_o) + e^{-U}$) (Charlesworth and Charlesworth 1997) seems to approximate the simulations reasonably well for a small equilibrium size of the best class (although this is highly dependent on the mutation rate), but greatly overestimates the time for a click when $n_o$ is large or $U$ is low (as seen in figures 2.4 - 2.7).
When $s$ is small one must account for the time to approach the deterministic equilibrium of the new least loaded class, which, in that range of $s$, becomes the major component of the total time to absorption.

The fact that, for a constant $n_0$ and constant $s$, no significant differences in the speed of the ratchet were observed, over an order of magnitude change in the population size (figure 2.1), suggests that $n_0s$ is an important parameter, although the expression for the average time between clicks of the ratchet is not an explicit function of $n_0s$. For the parameter range considered here, I found that an increase of 10-fold in $n_0s$ caused a decrease of around 10-fold in the speed of the ratchet. I also found that when $n_0s > 15$, the speed of the ratchet becomes exceedingly small. To support this, figure 2.12 shows the mean time as a function of $n_0s$ for a large number of simulations with very different values of $N$, $U$ and $s$.

![Figure 2.12](image)

**Figure 2.12** Dependence of the speed of the ratchet on $n_0s$. 

45
Because of the assumptions I have made to derive these results, some caution has to be taken considering their implications. First I must note that I have assumed, for simplicity, that all deleterious mutations have the same effect. However, recent work has suggested that an equal-effect of mutations assumption does not fit the data from mutation accumulation experiments, which are designed to measure the mutation rate to deleterious mutations and the selection coefficients against those mutations (Davies et al. 1999; Fry et al. 1999; Keightley 1996). The occurrence of many mutations with small effects, and a few with large effects, seems to be more consistent with the results. If this is the case, the ratchet is expected to turn at a much higher speed than for a single selection coefficient of the order of 1% (Gessler and Xu 1999; see chapter 3 and chapter 5), but each turn will cause a very small decline in mean fitness if many mutations have very low selection coefficients.

I have also assumed independence of mutational effects and it has been shown that epistasis slows down the speed of this process (Kondrashov 1994). However it has been shown that, if there is in fact a distribution of mutational effects with a more or less exponential shape, epistasis will not stop the ratchet (Butcher 1995).
3 The speed of Muller's ratchet with background selection, and the degeneration of Y chromosomes

This chapter is based on an article that will be published in *Genetical Research* (Gordo and Charlesworth in press)

3.1 Introduction

In the previous chapter, I have studied Muller's ratchet with the simplest possible mathematical model. I concluded that, if mutations cause sufficiently large deleterious effects, the expected time for them to accumulate in large populations is so long that they effectively do not accumulate by the ratchet. Such mutations are kept indefinitely at frequencies close to those expected under equilibrium between mutation and selection. However, their presence is known to interfere with the dynamics of linked neutral or very weakly selected mutations (those for which $N e s I \sim O(1)$), and even of mutations with much larger effects (Barton 1995; Charlesworth 1994; Charlesworth et al. 1993a; Fisher 1930; Peck 1994; Stephan et al. 1999). Their presence affects both nucleotide site diversity and fixation probabilities at weakly selected linked sites (Charlesworth 1994; Stephan et al. 1999), namely decreasing diversity, decreasing the probability of fixation of advantageous mutations, and increasing the probability of fixation of deleterious ones. In large populations with no recombination, this process can be quantified as though the effective population size is reduced approximately by a fraction $f_0$, the frequency of individuals free of such deleterious mutations. This is the process called background selection (Charlesworth et al. 1993a). The degeneration of large non-recombining portions of the genome such as the Y chromosome may be caused by the operation of Muller's ratchet (Charlesworth 1978) and/or the fixation of weak deleterious mutations due to background selection (Charlesworth 1996b). Because both strongly and mildly
deleterious mutations probably occur in natural populations (Keightley and Eyre-Walker 1999), it is important to quantify the effect of their simultaneous presence.

Here, the rate of accumulation due to the ratchet is analysed, using a simple model to study the effect of strongly deleterious mutations of effect $s_b$ (those that do not accumulate and can cause background selection) on the rate of accumulation of mildly deleterious mutations of effect $s_x$ (those that allow Muller’s ratchet to operate), in haploid asexual populations. It is shown that the rate of accumulation can be greatly increased by the presence of strongly deleterious mutations, and that the effect of the latter is generally to reduce the effective population size as predicted by the theory of background selection (Charlesworth et al. 1993a). Two approximations are proposed, that depend on the relative values of the parameters, to calculate the rate of accumulation of mildly deleterious mutations in this simple model. Simulation based studies are conducted for the more complex case where both types of deleterious mutations can accumulate. These are compared to the simple case of a population subject to just one type of mutation, whose effect is given by the harmonic mean of the effects in the two-type mutation model. As in the case of mutations with constant effects (Bergstrom and Pritchard 1998; Charlesworth and Charlesworth 1997; Higgs and Woodcock 1995), in this simple model the rate of accumulation is equal to the rate of fixation of deleterious mutations.

3.2 Theoretical considerations

3.2.1 Muller’s ratchet with background selection

Consider first an effectively infinite, non-recombining haploid population, subject to deleterious mutations whose effect on fitness is $s_x$, and which occur
according to a Poisson distribution with mean $U_s$. With multiplicative fitnesses, at mutation-selection balance, the expected frequency of individuals with $i$ mutations is given by a Poisson distribution with mean $\Theta_s = U_s/s_s$ (Haigh 1978; Kimura and Maruyama 1966).

Consider now that such a very large population is subject to deleterious mutations with two kinds of effect: $s_s$ and $s_b$ (with $s_b >> s_s$). Let the fitness of an individual carrying $k$ mutations with effect $s_s$ and $n$ mutations with effect $s_b$ be $(1-s_s)^k (1-s_b)^n$. Let $p_{nk}$ be the frequency of individuals carrying $k$ mutations with effect $s_s$ and $n$ mutations with effect $s_b$. The recursion for $p_{nk}$ is:

$$w p_{nk}(t+1) = \sum_{i=0}^{n} \sum_{j=0}^{k} p_{ij}(t)(1-s_s)^i (1-s_b)^j \frac{U_b^{n-i} U_s^{k-j}}{(n-i)! (k-j)!} e^{-U_b} e^{-U_s},$$

where $U_b$ and $U_s$ are the mutation rates for mutations with effect $s_b$ and $s_s$, respectively, and $\bar{w}$ is the population mean fitness.

Because new mutations are assumed to be Poisson-distributed, the recursion for $p_{00}$ implies that the mean fitness at equilibrium (c.f. Kimura and Maruyama 1966) is:

$$\bar{w} = e^{-U_b} = e^{-(U_b + U_s)}$$

By substitution, it is easy to see that an equilibrium solution is:

$$p_{0} = \frac{\Theta_b^i e^{-\Theta_b} \Theta_s^j e^{-\Theta_s}}{i! j!}$$

where $\Theta_b = U_b/s_b$ and $\Theta_s = U_s/s_s$ (which is a particular case of the general solution for a distribution of mutation effects obtained by Johnson (1999)). Although the stability of this equilibrium solution, has not been formally established, simulations suggest that it is a stable solution, when $N \to \infty$. 

49
The probability of having \( n = i+j \) mutations is given by the convolution of two Poisson distributions, which is a Poisson distribution with mean \( \Theta_b + \Theta_s \). At equilibrium the number of individuals free of mutations is thus:

\[
n_{00} = Ne^{-\Theta_s}e^{-\Theta_b}.
\]

Assume that \( N \) is large and that \( \Theta_b \) is small, so that strongly deleterious mutations are maintained close to equilibrium (which is expected to occur if \( N \exp(-\Theta_b) s_b > 15 \), as shown in the previous chapter). This implies that the fraction of the population that is free of these mutations is given by:

\[
f_{ob} = \exp(-U_b/s_b).
\]

The background selection principle (Charlesworth et al. 1993a) implies that neutral or weakly selected mutations can spread to fixation only if they arise on a background free of strongly deleterious ones (Barton 1995; Charlesworth 1994; Fisher 1930; Peck 1994). This suggests that mutation accumulation by the ratchet will only occur by an increase in the number of mildly deleterious mutations in a background free of strongly selected mutations, when these are being held at deterministic equilibrium.

This will affect the rate of the ratchet for mildly deleterious mutations as follows. Let \( y \) and \( z \) be the frequencies of the fittest class with respect to the mildly deleterious mutations of effect \( s_j \) and to the strongly deleterious of effect \( s_b \), so that \( x = y z \) is the frequency of the least loaded class, whose initial frequency is \( x_0 = \exp(-U_j/s_j)\exp(-U_b/s_b) \). I will determine the expected time until absorption (at \( x = 0 \)) using a diffusion equation as previously (Charlesworth and Charlesworth 1997; Stephan et al. 1993; and Chapter 2). Because, by assumption, \( z \) does not change, \( x = y \exp(-U_j/s_j)\exp(-U_b/s_b) \).
Therefore the diffusion equation for the density of time until absorption for \( x \) is identical to that for \( y \), but with a diffusion coefficient given by:

\[
b(y) = \frac{y}{N \exp \left( \frac{U_b}{s_b} \right)}
\]

assuming the Wright-Fisher model and \( y \ll 1 \).

Using the same assumptions as in chapter 2, the drift coefficient is approximated by:

\[
a(y) \approx 0.6s_y(1 - y/f_{0s})y
\]

where \( f_{0s} = \exp(-U/s_y) \). This means that the mean time to absorption (Ewens 1979, Chap. 4), i.e., the mean time for a click of the ratchet, is the same as in the case of a population subject to mutations with just one selection coefficient (of value \( s_y \)) occurring at a rate \( U_y \), but whose size is \( Nf_{0s} \), i.e. the size of the least-loaded class is discounted by a factor equal to the frequency of the least-loaded class with respect to the strongly selected mutations.

It is clear that, if the frequency of the class free of strong deleterious mutations is not constant, then a one-dimensional diffusion equation will not be adequate, and stochastic changes in the frequency of the strong class have to be taken into account, as well as the covariance between the frequency of the strong and mildly deleterious mutations. I was unable to find any way of dealing with this more complicated problem.

One must also account for the time taken for the value of the least-loaded class after a click (with respect to the mild effect mutations) to reach a value close to that expected at mutation-selection equilibrium, as has been done in section 2.5 of
the previous chapter. The total time for a click of Muller’s ratchet will then be given by:

\[
T(N,U_s,s_s,U_b,s_b) = \frac{1}{s_s} \left( 1 - 1.6 \frac{s_s}{U_s} \right) + \int_0^{2Nf_0b} \left( \int_0^x G(x')dx' \right) dx + \int_{f_0s}^{2Nf_0b} \left( \int_0^{f_0s} G(x')dx' \right) dx
\]

where

\[
G(x') = \exp \left( \frac{2Nf_0b}{f_0s} 0.6s_s \left( \frac{x'}{2} - f_0s \right) \right).
\]

The first term represents the approximate time to approach a value close to the equilibrium and the other two terms represent the mean time to absorption of the stochastic process with initial frequency \( f_0s \).

### 3.3 Simulation methods

For a given population size \( N \), genomic mutation rate to deleterious mutations \( U \), and selection coefficients against each mutation \( s_s \) or \( s_b \), haploid asexual populations were simulated, starting with all individuals initially free of mutations. Assuming the same sequence of events as in the previous chapter, populations were then run for an initial period of generations so that they could approach an equilibrium between mutation and selection (i.e. the mean fitness of the population reaches the value \( \exp(-U) \)). An additional time until a click of the ratchet for the mutations of small effect was observed was allowed. After this initial period, populations were run for more than 2,000 generations, and up to 100,000 generations, depending on the speed at which mutations accumulated. Every generation, the number of mildly selected mutations and strongly selected mutations in every individual were counted, and the number of individuals with the smallest number of mutations (the least-loaded class) recorded. If, at a given
generation, the number of mutations in the least-loaded class increases, the ratchet has clicked.

To form a new generation, individuals are sampled randomly from the previous generation, then subjected to the occurrence of mutations sampled from a Poisson distribution, that can have either an effect $s_s$ (mildly deleterious mutations) or $s_b$ (strongly deleterious mutations) on fitness, and assigned probabilities of survival $(1-s_s)^k(1-s_b)^n$, where $k$ is the number of mutations with effect $s_s$ and $n$ is the number of mutations with effect $s_b$ carried by an individual. A new generation of $N$ individuals is constructed by comparing the probability of survival of each individual with a pseudo-random number drawn from a uniform distribution in the interval [0, 1]. Each run was repeated 5 times and the average rate of accumulation of mutations, measured by the average time to loss of the least-loaded class, was computed.

To check the relation between losses of the least-loaded class and fixation events, a number of multi-locus stochastic simulations were performed, in which the fate of each mutation at every locus was followed and the fixation events counted.

3.4 Results

3.4.1 The rate of accumulation by the ratchet with background selection

3.4.1.1 Region of parameter space for which $n_{00} > 1$

As discussed above, if the rate at which strong mutations occur ($U_b$) is not too large, their effect is not too small, and $N$ is large, they will stay near their expected equilibrium frequency, and the rate of accumulation of the mildly deleterious is the same as in a population whose size is reduced by the factor $f_{ob}$ (Charlesworth et al. 1993a).
In figures 3.1 and 3.2 I plot the average time to successive losses of the least-loaded class measured in the simulations, with a population subject to mutations with two types of effect (T2-class model) and for the case where we have a population of reduced size, $f_{ob}$, subject to mutations of just one type of effect ($s=s_b$), occurring at a rate $U_b$ (T1-class model). I also plot the theoretical prediction based on equation (1) (Theoretical). In figure 3.1 the time between clicks is shown as an increasing function of the population size ($N$), and in figure 3.2 as a function of the effects of the weaker mutations ($s_j$) in a population of $10^5$ individuals. For the case when $s_j = 0.002$, the theoretical approximation is invalid because $n_{\infty} < 1$, and another expression should be used to predict the rate of mutation accumulation (Gessler 1995), as described below.

![Figure 3.1](image-url)

**Figure 3.1** The mean time between clicks of the ratchet as a function of the population size $N$ in the presence of background selection (T2-class model). The mutation rate for mildly deleterious mutations with selection coefficient $s = 0.005$ is $U_s = 0.03$, and the rate for strongly deleterious mutations with selection coefficient $s_b = 0.02$ is $U_b = 0.01$. The T1-class model is the mean time from simulations with populations of size $f_{ob}N = N \exp(-U_b/s_b)$, mutation rate $U_s$ and selection coefficient $s_j$.

Theoretical is the time calculated from equation (3).
Figure 3.2 The average time for a click of the ratchet, as a function of the effect of mildly deleterious mutations, in a population of $10^5$ individuals. $s_b = 0.02$, $U_b = 0.01$ and $U = 0.03$. Other symbols as in figure 3.1. For $s = 0.002$, $n_{\infty}$ is less than 1 and the diffusion approximation is invalid.

Figures 3.1 and 3.2 mostly display cases where the expected size of the least-loaded class in an effectively infinite population, $n_{\infty}$, is greater than 1. It is clear that, under this condition, the rate of accumulation of deleterious mutations is increased when compared with the case of occurrence of mutations with just one effect, due to the reduction in $N_e$ and that it can be predicted with reasonable accuracy using equation (3). Several other parameter sets were simulated that support this conclusion (results not shown). Therefore, even large populations can suffer from a very rapid accumulation of mildly deleterious mutations if the reduction in $N_e$ due to strong deleterious mutations is large. To better see the strength of this effect on the speed of the ratchet, Table 3.1 shows the effect of background selection on the average time between clicks of the ratchet as found in the simulations. Note that, in all cases in Table 3.1, no accumulation of strongly deleterious mutations occurs.
Table 3.1.
The time between clicks of Muller's ratchet ($T$) in the absence and presence of background selection

<table>
<thead>
<tr>
<th>Presence of background selection</th>
<th>Absence of background selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U_b$ = 0.01 $s_b$ = 0.02</td>
<td>$U_s$ = 0.03 $s_s$ = 0.005</td>
</tr>
<tr>
<td>$U_b = 0.01$ $s_b = 0.01$</td>
<td>$U_s = 0.03$ $s_s = 0.005$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$N$</th>
<th>$T$</th>
<th>2SE</th>
<th>$T$</th>
<th>2SE</th>
<th>$T$</th>
<th>2SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>5000</td>
<td>228</td>
<td>12</td>
<td>185</td>
<td>13</td>
<td>182</td>
<td>27</td>
</tr>
<tr>
<td>10000</td>
<td>324</td>
<td>32</td>
<td>256</td>
<td>18</td>
<td>229</td>
<td>11</td>
</tr>
<tr>
<td>50000</td>
<td>778</td>
<td>108</td>
<td>579</td>
<td>85</td>
<td>413</td>
<td>26</td>
</tr>
<tr>
<td>100000</td>
<td>1144</td>
<td>175</td>
<td>875</td>
<td>80</td>
<td>649</td>
<td>79</td>
</tr>
<tr>
<td>500000</td>
<td>12819</td>
<td>5681</td>
<td>4114</td>
<td>551</td>
<td>2875</td>
<td>1189</td>
</tr>
</tbody>
</table>

Results on the effect of background selection, for different values of population size ($N$) and mutation rates to weakly and strongly deleterious mutations, $U_s$ and $U_b$, respectively. The corresponding selection coefficients are $s_s$ and $s_b$. SE is the standard error. Note that no accumulation of the strongly deleterious mutations was observed in the simulations.

Although the rate of accumulation of mildly deleterious mutations in the presence of strong ones is approximately the same as for a population whose size is reduced by the factor $f_{ob}$ and which has the mutation rate $U_s$, one observes that the rate of accumulation under background selection seems to be systematically lower than expected, although the difference is relatively small. For example, in figure 3.2, the time for the T2-class model is 13,965 generations but that for the T1-class model is 10,074 for $s_s = 0.007$. In fact, if one follows the value of the fraction of individuals that are free of strong deleterious mutations, one observes that the average value of this fraction is slightly bigger than the expected value of $f_{ob}$. Therefore, if this is always true, the reduction in effective population size is slightly smaller than that predicted by the deterministic equilibrium value, so the diffusion coefficient in the diffusion approximation is smaller, which will increase the time to absorption.
3.4.1.2 Region of parameter space for which \( n_{00} < 1 \)

In the above analysis, I have quantified the ratchet mechanism as classically defined by Muller (1964, p.2), i.e. loss of the least-loaded class due to drift after the equilibrium between mutation and selection is reached. This requires \( N_{S_b} \gg 1, N_{S_s} \gg 1 \) and \( n_{00} > 1 \). When \( N_{S_b} \gg 1 \) and \( N_{S_s} \gg 1 \) but \( n_{00} < 1 \), mutation accumulation is a quasi-deterministic process and stochasticity plays a second-order role in driving the ratchet (Gessler 1995). Under these circumstances, Gessler has proposed an expression for calculating the rate of the ratchet (equation (8) in Gessler (1995), based on parameterisation of the distribution of the number of mutations in the population.

![Figure 3.3](image)

**Figure 3.3** The average time for a click of the ratchet as a function of the population size for the range of parameter space where \( n_{00} < 1 \). Theoretical is the theoretical value based on equation (8) in Gessler (1995), adjusted for reduced \( N_e \). Other symbols as in figure 3.1. \( U_b = 0.05 \), \( s_b = 0.05 \), \( U_s = 0.15 \), \( s_s = 0.01 \).

In figure 3.3, I show the rate of accumulation for some cases where these conditions are met. Even in these cases, the value for the time is reasonably well predicted by the reduction in population size \( f_{ob} \). The theoretical value (Theoretical) obtained with
a reduced population size under the approximation of Gessler (1995) seems to predict the rate of mutation accumulation quite well, for this region of parameter space.

3.4.2 Dynamics of very weakly deleterious mutations

When mutations have extremely weak effects, such that $N_s$ or $f_{ob}N_s$, is of the order of 1, their dynamics are mainly determined by drift. This means that the frequencies of these mutations can reach high values, and the assumption of negligible back-mutation becomes unrealistic. This implies that the most important component of the ratchet mechanism, its irreversibility, is likely to be lost in real biological systems. Studies of completely linked sets of weakly selected loci with reversible mutation have been done by Li (1987), Comeron et al. (1999), McVean and Charlesworth (2000) and Tachida (2000).

One point to note is the fact that, when $f_{ob}N_s >> 1$, the rate of the ratchet is much bigger than the rate of fixation of deleterious mutations due to drift calculated from the standard single-locus formula (Crow and Kimura 1970, chap. 8; Higgs and Woodcock 1995). For a given $N$ and $s$, the rate of fixation due to drift increases linearly with $U$, but the rate of the ratchet increases in an exponential fashion (Charlesworth et al. 1993b; see also figure 2.7 in the previous chapter). This behaviour of the ratchet is not seen when $N_s$, or $f_{ob}N_s < 1$, as shown in Table 3.2.

Clearly, when $f_{ob}N_s > 1$ the rate of the ratchet is greater than the rate of fixation due to drift, but, as this quantity tends to or becomes smaller than 1, the rate of the ratchet and the single-locus rate of fixation converge to the same value, approaching the mutation rate $U$, from below as $f_{ob}N_s \rightarrow 0$. For $f_{ob}N_s << 1$, this means that mutation accumulation is mainly occurring due to independent fixations of deleterious mutations at different loci. It is therefore not correct to call this a
ratchet, since there is no acceleration of the rate of fixation over that expected with independence between loci.

Table 3.2.
Comparison of the rate of fixation of weak deleterious mutations due to drift ($R_{\text{fix}}$) with the rate of the ratchet ($R_{\text{rat}}$).

<table>
<thead>
<tr>
<th>$N$</th>
<th>$U_b$</th>
<th>$s_b$</th>
<th>$U_s$</th>
<th>$s_s$</th>
<th>$n_{00}$</th>
<th>$N$</th>
<th>$s_f$</th>
<th>$f_{00}N$</th>
<th>$R_{\text{rat}}$</th>
<th>$R_{\text{fix}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.0025</td>
<td>7.6x10^{-7}</td>
<td>2.5</td>
<td>0.92</td>
<td>0.0025</td>
<td>0.03</td>
<td>0.02</td>
</tr>
<tr>
<td>0.05</td>
<td>0.005</td>
<td>1.7x10^{-2}</td>
<td>5</td>
<td>1.84</td>
<td>0.02</td>
<td>0.005</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td>0.01</td>
<td>2.5</td>
<td>10</td>
<td>3.68</td>
<td>0.015</td>
<td>0.0002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td>0.02</td>
<td>30</td>
<td>20</td>
<td>7.36</td>
<td>0.005</td>
<td>3 x10^{-7}</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40000</td>
<td>0.028</td>
<td>0.042</td>
<td>0.0001</td>
<td>9.6x10^{-180}</td>
<td>4</td>
<td>0.24</td>
<td>0.035</td>
<td>0.033</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>0.0002</td>
<td>1.5x10^{-88}</td>
<td>8</td>
<td>0.49</td>
<td>0.03</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>0.0005</td>
<td>8.0x10^{-34}</td>
<td>20</td>
<td>1.22</td>
<td>0.03</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100000</td>
<td>0.035</td>
<td>0.015</td>
<td>0.0001</td>
<td>6.5x10^{-64}</td>
<td>10</td>
<td>0.01</td>
<td>0.014</td>
<td>0.015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.006</td>
<td>0.0001</td>
<td>2.1x10^{-63}</td>
<td>10</td>
<td>0.03</td>
<td>0.014</td>
<td>0.015</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>0.0001</td>
<td>2.2x10^{-62}</td>
<td>10</td>
<td>0.30</td>
<td>0.012</td>
<td>0.011</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.03</td>
<td>0.0001</td>
<td>2.2x10^{-61}</td>
<td>10</td>
<td>3.11</td>
<td>0.01</td>
<td>0.0002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$n_{00}$ is the equilibrium size of the least-loaded class. Other parameters as in Table 1. The rate of fixation is calculated using the standard single-locus formula for the rate of fixation of deleterious mutations: $R_{\text{fix}} = f_{00}NU_f \frac{\exp(2s_f) - 1}{\exp(2f_{00}Ns_f) - 1}$ (Crow and Kimura, Chap. 8, 1970)

In all the cases shown in Table 3.2, $N_s >> 1$, so that the rate of fixation due to drift of the small effect mutations would be negligible if they occurred in isolation. It is the effect of background selection that allows them to experience an accelerated rate of fixation (Charlesworth 1994; Charlesworth 1996b; Charlesworth and Charlesworth 1998). Even in the presence of a ratchet for the large effect mutations, for the cases where $N = 10^5$ and $s_b$=0.005 or 0.006 in Table 3.2, the rate of fixation of the small effect mutations seems to be well predicted by a simple reduction in $N$, by a factor of $f_{00}$. This is because the distribution of the large effect mutations is close to the deterministic equilibrium expectation when $N$ is large (Charlesworth and
Charlesworth 1997), so that the size of the class free of such mutations stays close to $f_{0b}N$ over long periods of time (as observed in the simulations).

![Figure 3.4](Image.png)

**Figure 3.4.** The change in mean fitness over time with reversible mutation occurring at a rate $U_{\text{back}}$ per individual per generation. $N = 10,000$, $U_b = 0.05$, $s_b = 0.02$, $U_\text{is} = 0.05$, $s_\text{is} = 0.001$. $U_{\text{back}}$ is 0 for triangles, 0.01 for circles, 0.02 for squares and 0.03 for diamonds. The back-mutation rate is with respect to the weakly selected mutations.

In figure 3.4, the effect of back-mutation is considered for cases where $f_{0b}N_s < 1$. Back-mutation only occurs for the weakly selected mutations. Like it is intuitively obvious, the bigger the back-mutation rate ($U_{\text{back}}$), the smaller the decline in log mean fitness. Note that, when there is back-mutation, one can no longer talk about a ratchet because, even if the least-loaded class gets lost at some point in time, this loss is not irreversible, since it can be rebuilt, and an equilibrium is eventually established even with no recombination (McVean and Charlesworth 2000).

**3.4.3 Accumulation of both types of mutations**

In addition to the accumulation of the small-effect mutations, stochastic loss of the least-loaded class of the large-effect mutations may also occur. This is
especially likely in very small populations, when the mutation rate to the large-effect mutations is large or when the difference in the selection coefficients is not very large. In Table 3.3, I show the results of simulations for some cases where this occurs. I also show the rate of decline in log mean fitness per generation due to the accumulation of both types of mutations ($\Delta \ln \bar{w}$). I then compare this decline with the one occurring in a population subject to mutations whose fitness effect is equal to the harmonic mean ($s_H$) of the fitness effects of the corresponding two-type mutation model:

$$\frac{1}{s_H} = \frac{1}{(U_p + U_s)} \left( \frac{U_p}{s_p} + \frac{U_s}{s_s} \right)$$

This reflects the fact that the mean number of mutations, and the deterministic equilibrium size of the least-loaded class, are the same in both cases. I find that, in these small populations, when the rates of accumulation of each type of mutation are of the same order of magnitude, the rate of decline in log mean fitness seems to be well approximated by using the harmonic mean of the selection coefficients.

If one equates the decline in log mean fitness in both cases we will have:

$$- \left[ \frac{s_s}{T(s_s)} + \frac{s_b}{T(s_b)} \right] \approx - \left[ \frac{s_H}{T(s_H)} \right]$$

where $T(s_s)$ and $T(s_b)$ are the times to loss of the least-loaded class for each mutation type.

Let $K_s$ and $K_b$ be the probabilities that a click of the ratchet involves the class of small and large effect mutations, respectively. We can rewrite the above expression as follows:

$$- \left[ \frac{s_s}{T(s_H)} + \frac{s_b}{T(s_H)} \frac{K_s}{T(s_H)} + \frac{K_b}{T(s_H)} \right] \approx - \frac{1}{\left( \frac{U_p}{s_p} + \frac{U_s}{s_s} \right)} \left[ \frac{U_s}{T(s_H)} + \frac{U_b}{T(s_H)} \right]$$
Then one can, tentatively, approximate the probabilities by:

\[
K_s = \frac{U_s}{s_s + U_b/s_b} \quad \text{and} \quad K_b = \frac{s_b}{s_s + U_b/s_b}
\]

(4)

i.e. the proportion of the mean number of mutations of each class.

Table 3.3.

The relation between a mutation model with two types of mutational effects and a model with a constant mutation effect equal to the harmonic mean of the two effects (s_H).

<table>
<thead>
<tr>
<th>N</th>
<th>U_s</th>
<th>s_s</th>
<th>U_b</th>
<th>s_b</th>
<th>s_H</th>
<th>T(s_s)</th>
<th>T(s_b)</th>
<th>T(s_H)</th>
<th>Δln w</th>
<th>Δln w_{s_H}</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>0.05</td>
<td>0.02</td>
<td>0.05</td>
<td>0.03</td>
<td>0.024</td>
<td>90</td>
<td>222</td>
<td>70</td>
<td>-3.6x10^4</td>
<td>-3.4x10^4</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.01</td>
<td>0.05</td>
<td>0.03</td>
<td>0.015</td>
<td>52</td>
<td>379</td>
<td>40</td>
<td>-2.7x10^4</td>
<td>-2.8x10^4</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.01</td>
<td>0.05</td>
<td>0.015</td>
<td>0.012</td>
<td>57</td>
<td>92</td>
<td>33</td>
<td>-3.4x10^4</td>
<td>-3.6x10^4</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.0075</td>
<td>0.05</td>
<td>0.015</td>
<td>0.01</td>
<td>44</td>
<td>102</td>
<td>32</td>
<td>-3.2x10^4</td>
<td>-3.1x10^4</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.01</td>
<td>0.075</td>
<td>0.015</td>
<td>0.013</td>
<td>50</td>
<td>55</td>
<td>25</td>
<td>-4.7x10^4</td>
<td>-4.9x10^4</td>
</tr>
<tr>
<td></td>
<td>0.025</td>
<td>0.02</td>
<td>0.075</td>
<td>0.03</td>
<td>0.027</td>
<td>164</td>
<td>145</td>
<td>85</td>
<td>-3.3x10^4</td>
<td>-3.4x10^4</td>
</tr>
<tr>
<td>1000</td>
<td>0.05</td>
<td>0.02</td>
<td>0.05</td>
<td>0.03</td>
<td>0.024</td>
<td>141</td>
<td>389</td>
<td>97</td>
<td>-2.2x10^4</td>
<td>-2.5x10^4</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.01</td>
<td>0.05</td>
<td>0.03</td>
<td>0.015</td>
<td>57</td>
<td>898</td>
<td>49</td>
<td>-2.1x10^4</td>
<td>-3.1x10^4</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.0075</td>
<td>0.05</td>
<td>0.015</td>
<td>0.012</td>
<td>46</td>
<td>1199</td>
<td>47</td>
<td>-1.9x10^4</td>
<td>-2.6x10^4</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.01</td>
<td>0.075</td>
<td>0.015</td>
<td>0.013</td>
<td>58</td>
<td>69</td>
<td>31</td>
<td>-3.9x10^4</td>
<td>-4.0x10^4</td>
</tr>
<tr>
<td></td>
<td>0.025</td>
<td>0.02</td>
<td>0.075</td>
<td>0.03</td>
<td>0.027</td>
<td>252</td>
<td>277</td>
<td>152</td>
<td>-1.9x10^4</td>
<td>-1.8x10^4</td>
</tr>
</tbody>
</table>

T(s_s) and T(s_b) are the times to loss of the least-loaded class of mutations with selection coefficient s_s and s_b, respectively. T(s_H) is the time for a click of the ratchet in a population subject to mutations with fitness effect s_H. Δln w and Δln w_{s_H} are the rates of decline in mean fitness for each mutational model. Other parameters as in Table 3.1.

Since T(s_H) can be calculated, with these expressions I can calculate the rate of movement for each class. As illustrated in Table 3.4, this approximation gives accurate results only in cases where the mutation rates and selection coefficients for each class are fairly similar.
Table 3.4.
The times to loss of each class when there is accumulation of both types of mutations.

<table>
<thead>
<tr>
<th>$U_s$</th>
<th>$s_s$</th>
<th>$U_b$</th>
<th>$s_b$</th>
<th>$T(s_s)$</th>
<th>$T(s_b)$</th>
<th>$T(s_s)$</th>
<th>$T(s_b)$</th>
<th>$\Delta ln w$</th>
<th>$\Delta ln w_{s_b}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.04</td>
<td>0.001</td>
<td>0.04</td>
<td>0.01</td>
<td>34</td>
<td>835</td>
<td>34</td>
<td>341</td>
<td>$4 \times 10^{-5}$</td>
<td>$6 \times 10^{-5}$</td>
</tr>
<tr>
<td>0.04</td>
<td>0.002</td>
<td>0.04</td>
<td>0.01</td>
<td>43</td>
<td>455</td>
<td>47</td>
<td>234</td>
<td>$7 \times 10^{-5}$</td>
<td>$8 \times 10^{-5}$</td>
</tr>
<tr>
<td>0.04</td>
<td>0.005</td>
<td>0.04</td>
<td>0.01</td>
<td>90</td>
<td>309</td>
<td>100</td>
<td>201</td>
<td>$9 \times 10^{-5}$</td>
<td>$1 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.04</td>
<td>0.009</td>
<td>0.04</td>
<td>0.01</td>
<td>181</td>
<td>228</td>
<td>205</td>
<td>228</td>
<td>$9 \times 10^{-5}$</td>
<td>$9 \times 10^{-5}$</td>
</tr>
<tr>
<td>0.04</td>
<td>0.005</td>
<td>0.06</td>
<td>0.0075</td>
<td>93</td>
<td>86</td>
<td>88</td>
<td>88</td>
<td>$1 \times 10^{-4}$</td>
<td>$1 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.02</td>
<td>0.005</td>
<td>0.04</td>
<td>0.008</td>
<td>268</td>
<td>234</td>
<td>236</td>
<td>189</td>
<td>$5 \times 10^{-5}$</td>
<td>$6 \times 10^{-5}$</td>
</tr>
<tr>
<td>0.01</td>
<td>0.0025</td>
<td>0.02</td>
<td>0.008</td>
<td>328</td>
<td>2699</td>
<td>483</td>
<td>772</td>
<td>$1 \times 10^{-5}$</td>
<td>$2 \times 10^{-5}$</td>
</tr>
</tbody>
</table>

The population size is 10,000. The predicted times are calculated using equations (4a) and (4b). Other parameters as in Table 3.3. $\Delta ln w$ and $\Delta ln w_{s_b}$ are the declines in mean fitness from the simulations.

From many different parameters values tested, for which accumulation of both mutations occurs, I draw the following conclusion: as long as the rates of movement of each class are of the same order of magnitude, the decline in log mean fitness seems to be well approximated using the harmonic mean of the selection coefficients. When this is true, and when the mutation rates and selection coefficients for each class of mutations are similar, the rate of movement of each class seems to be reasonably well approximated using equations (4). When one of the mutational classes starts to accumulate at a rate that is an order of magnitude lower than the other, and $s_b >> s_s$, one can use the reduction in effective population approximation to predict the rate of movement of the latter, and approximate the rate of movement of the large-effect deleterious class by treating it as though the small-effect mutations were absent.
3.5 Discussion

For mathematical simplicity the quantification of the rate of the ratchet is usually done under the assumption of equal mutation effects across loci. Although there have been some studies in which a continuous distribution of mutation effects was considered (Butcher 1995; Gessler 1995), analytical expressions for the rate of this process exist only for cases of constant selection coefficients. Recent studies on the rate of occurrence of deleterious mutations and on the values of the selection coefficients of those mutations have cast doubts on the simple assumption that all detrimental mutations have similar deleterious effects (Davies et al. 1999; Keightley 1996). Davies et al. (1999) have studied the effects of induced mutations in *C. elegans* and observed that the distribution of mutation effects appears to be bimodal. A simple discrete distribution with two classes of mutation effects seems to fit the data better than a continuous gamma distribution. While mutations with mean heterozygous fitness effects of the order of 1-2% have been inferred from laboratory experiments on *Drosophila melanogaster* (Charlesworth and Hughes 2000), there may well exist a numerous class of mutations with much smaller effects (Davies et al. 1999), which are simply very difficult to detect in classic mutation accumulation experiments. I have tried to quantify the process of mutation accumulation by Muller's ratchet in these circumstances.

3.5.1 The rate of the ratchet under background selection

With a simple two-category distribution of mutation effects (\(s_b\) and \(s_s\)), in which there is a strong asymmetry in mutation rates (\(U_b < U_s\)), namely those mutations that are more deleterious (\(s_b\)) occur at a lower rate than those that are less deleterious (\(s_s\)), I found that the effect of the strong mutations, which do not
accumulate, on the rate of accumulation of the weaker ones can be generally well predicted, in the absence of recombination, by the background selection principle with the appropriate reduction in effective population size (Charlesworth et al. 1993a; Fisher 1930). As with the equal-effects mutation model, one can use an explicit expression to quantify the rate of the ratchet, under conditions when the large-effect mutations do not accumulate.

3.5.2 The rate of the ratchet in small populations

In sufficiently small populations, both large and small effect mutations are likely to accumulate at high rates. In these circumstances, I have compared the decline in log mean fitness of the population with the decline in log mean fitness of a population subject to mutations whose effect $s_{ij}$ is given by the harmonic mean of the selection coefficients of the two classes of mutations (Gessler 1995). If both mutations accumulate at similar rates, then the rate of decline in log mean fitness is well approximated by the harmonic mean formula (as shown in Table 3.4). But if one of the mutation types accumulates at a much lower rate than the other type, then the background selection effect becomes evident, and this approximation is no longer appropriate (as shown in Tables 3.3 and 3.4).

In a haploid asexual population subject to deleterious mutations of equal effects, it is known that there is a one-to-one correspondence between the loss of the least-loaded classes and the fixation of deleterious mutations in the entire population (Bergstrom and Pritchard 1998; Charlesworth and Charlesworth 1997; Higgs and Woodcock 1995). In order to check the rate of advance of the ratchet with the rate of fixation events in the two-type mutation model, I ran multi-locus simulations where the fate of each mutation was followed. I observed that fixation events follow clicks of the ratchet, as previously thought.
3.5.3 Y and neo-Y chromosomes

As briefly mentioned in the introduction, Muller’s ratchet and/or background selection may have played a role in sex chromosome differentiation in animals (Charlesworth 1978; Charlesworth 1996b) and plants (Filatov et al. 2000), in which sex is determined by a system of X and Y chromosomes. Systems like this have evolved repeatedly and independently throughout evolutionary history. Sex chromosomes are thought to have evolved from a set of homologous, proto-X and proto-Y chromosomes, which stopped recombining completely or over a very large region. For example, there is some evidence that the evolution of the mammalian Y chromosome has been punctuated by at least 4 events that suppressed recombination between the X and the Y (Lahn and Page 1999), the first event having occurred about 300 million years ago.

The Y chromosome of the plant Silene latifolia, and the neo-Y chromosome of Drosophila miranda, which results from a fusion between a chromosome arm and the old Y chromosome, are much more recent than this (Yi and Charlesworth 2000; Bachtrog and Charlesworth 2000; Filatov et al. 1999). The same principles that apply to a primeval non-recombining Y chromosome also apply to a neo-Y chromosome of Drosophila. The erosion of a proto-Y or neo-Y chromosome is very similar to the degeneration of a haploid asexual population if one replaces $s$ by $hs$, where $h$ is the dominance coefficient and $s$ the selection coefficient effect against homozygous mutations (Charlesworth and Charlesworth 1997). One can therefore use the results just presented to try and investigate the role of this processes in different systems.

One very interesting system that has recently received some attention is the neo-Y chromosome system of Drosophila miranda (Charlesworth and Charlesworth 2000; Rice 1996; Steinemann and Steinemann 1992). It constitutes an excellent clock to set the time scale over which the degeneration of a non-recombining region
is supposed to occur. The time estimated for the origin of the chromosomal rearrangement generating the neo-\(Y\) in \textit{Drosophila miranda} is about 1.25 Mya (Yi 2000). The neo-\(Y\) shows evidence for degeneration, and the neo-\(X\) is partially dosage compensated (Steinemann and Steinemann 1998; Steinemann et al. 1993). These observations seem to suggest that, if there is a general process responsible for the degeneration of this non-recombining segment of the genome such as Muller’s ratchet, it is expected to show its signature over a time scale of the order of \(10^6\) years.

While estimates of the deleterious mutation rate and selection coefficients against these mutations exist from various studies on mutation accumulation in \textit{Drosophila}, the values of these parameters for mammals and plants are not so well known (Keightley and Eyre-Walker 1999). This makes the quantification of the combined role of Muller’s ratchet and background selection somewhat difficult. Tentatively, one can ask about the approximate time scale at which these two processes are expected to occur in these different organisms.

The effective population size of \textit{D. miranda} lies between \(10^5\) and \(10^6\) (Bachtrog and Charlesworth 2000; Yi and Charlesworth 2000). Its neo-\(Y\) chromosome constitutes about \(1/5\) of the genome. Estimates of the haploid genomic deleterious mutation rate in \textit{Drosophila} vary between 0.5 to 0.05 (Table 1 of Keightley and Eyre-Walker 1999). Assume, for the sake of argument, that the mutation rate for the neo-\(Y\) to strongly deleterious mutations, which can be detected in mutation accumulation experiments (i.e. with a heterozygous selection coefficient of 1-2\%), is 0.01. The haploid effective population size after correction due to background selection exerted by the strong mutations will then lie between \(2\times10^5\) and \(3\times10^5\) (taking account of the fact that the number of breeding males is half the population size). Every mutation whose effect is \(< 0.5\times10^{-5}\) will be effectively neutral in the presence of these strongly deleterious mutations and will experience an accelerated rate of fixation (Charlesworth 1996b). For mutations whose effect is
bigger than this, the ratchet mechanism will take place under background selection. In Table 3.5, I quantify the expected rate of the ratchet in the presence of background selection for mutations with different selection coefficients. This quantification is done, as explained previously, using equation (3) when \( n_{oo} > 1 \), and equation (8) of Gessler (1995), using a reduced effective population size, when \( n_{oo} < 1 \). I also show the time scale of this process for the Y chromosome of the plant *Silene latifolia*, assuming an effective population size of around \( 10^6 \) (Filatov et al. 2000). The values for the Y chromosome deleterious mutation rate assumed here are about one-twelfth the haploid deleterious mutation rate estimates from some species of selfing plants (Drake et al. 1998). This uses the number of chromosomes in *Silene latifolia* as a very rough guide, despite the fact that the X is bigger than the autosomes and the Y bigger than the X (Matsunaga et al. 1999). The large selection coefficients considered here are lower than the estimated values from the mutation accumulation experiments in *Arabidopsis thaliana* (Schultz et al. 1999), since these are probably overestimates.

The effective population size of mammals is probably at least an order of magnitude smaller than that of *Drosophila* or *Silene* (Charlesworth 1996b). Given that the deleterious mutation rate, at least in humans, is very high (Eyre-Walker and Keightley 1999; Nachman and Crowell 2000), if the first inversion to occur on the mammalian Y chromosome (Lahn and Page 1999) caused a region of about half of its size to stop recombining, then the reduction in effective population size due to background selection might have caused a rapid accumulation and fixation of mildly deleterious mutations due to the ratchet. Given that the X chromosome of eutherian mammals represents 3 to 5% of the haploid genome (Deloukas et al. 1998; Graves 1995), a deleterious mutation rate for the proto-Y of 0.04 can be considered reasonable, from current estimates for the deleterious mutation rate in hominids (Eyre-Walker and Keightley 1999).
Table 3.5.

Mean fitness and expected number of fixations after 1 million generations

<table>
<thead>
<tr>
<th>N</th>
<th>$u_i$</th>
<th>$s_i$</th>
<th>$u_b$</th>
<th>$s_b$</th>
<th>$n_{oo}$</th>
<th>$T$ (generations)</th>
<th>$w/w_i$</th>
<th>No. of fixations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Drosophila miranda</strong> neo-Y chromosome</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>125000</td>
<td>0.024</td>
<td>0.005</td>
<td>0.008</td>
<td>0.02</td>
<td>690</td>
<td>4508 (^a)</td>
<td>0.31</td>
<td>222</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(10446)</td>
<td>(0.62)</td>
<td>(96)</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.02</td>
<td>3x10^6</td>
<td></td>
<td></td>
<td>250 (^b)</td>
<td>0.018</td>
<td>4000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(333)</td>
<td>(0.05)</td>
<td>(3003)</td>
</tr>
<tr>
<td></td>
<td>0.005</td>
<td>0.01</td>
<td>462</td>
<td></td>
<td></td>
<td>2440 (^a)</td>
<td>0.13</td>
<td>410</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(10446)</td>
<td>(0.62)</td>
<td>(96)</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.01</td>
<td>2x10^6</td>
<td></td>
<td></td>
<td>250 (^b)</td>
<td>0.018</td>
<td>4000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(333)</td>
<td>(0.05)</td>
<td>(3003)</td>
</tr>
<tr>
<td></td>
<td>0.0005</td>
<td>0.01</td>
<td>8x10^{-17}</td>
<td></td>
<td></td>
<td>133 (^b)</td>
<td>0.023</td>
<td>7519</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(133)</td>
<td>(0.023)</td>
<td>(7519)</td>
</tr>
<tr>
<td></td>
<td>0.005</td>
<td>0.016</td>
<td>0.02</td>
<td>462</td>
<td></td>
<td>2440 (^a)</td>
<td>0.13</td>
<td>410</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(10446)</td>
<td>(0.62)</td>
<td>(96)</td>
</tr>
<tr>
<td><strong>Silene latifolia</strong> Y chromosome</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500000</td>
<td>0.0075</td>
<td>0.002</td>
<td>0.0025</td>
<td>0.01</td>
<td>9158</td>
<td>3x10^7 (^a)</td>
<td>1.00</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(6x10^8)</td>
<td>(1.00)</td>
<td>(0)</td>
</tr>
<tr>
<td></td>
<td>0.003</td>
<td>0.01</td>
<td>5x10^6</td>
<td></td>
<td></td>
<td>1111 (^b)</td>
<td>0.76</td>
<td>900</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1111)</td>
<td>(0.76)</td>
<td>(900)</td>
</tr>
<tr>
<td></td>
<td>0.015</td>
<td>0.002</td>
<td>0.005</td>
<td>0.02</td>
<td>215</td>
<td>1306 (^a)</td>
<td>0.22</td>
<td>766</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1548)</td>
<td>(0.27)</td>
<td>(646)</td>
</tr>
<tr>
<td></td>
<td>0.0005</td>
<td>0.02</td>
<td>4x10^8</td>
<td></td>
<td></td>
<td>400 (^b)</td>
<td>0.29</td>
<td>2500</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(400)</td>
<td>(0.29)</td>
<td>(2500)</td>
</tr>
<tr>
<td><strong>Mammalian Y chromosome</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50000</td>
<td>0.03</td>
<td>0.007</td>
<td>0.01</td>
<td>0.02</td>
<td>417</td>
<td>2442 (^a)</td>
<td>0.056</td>
<td>410</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(6356)</td>
<td>(0.33)</td>
<td>(157)</td>
</tr>
<tr>
<td></td>
<td>0.005</td>
<td>0.02</td>
<td>75</td>
<td></td>
<td></td>
<td>469 (^a)</td>
<td>2x10^{-3}</td>
<td>2132</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(663)</td>
<td>(5x10^{-4})</td>
<td>(1508)</td>
</tr>
<tr>
<td></td>
<td>0.005</td>
<td>0.06</td>
<td>105</td>
<td></td>
<td></td>
<td>586 (^a)</td>
<td>2x10^{-4}</td>
<td>1706</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(663)</td>
<td>(5x10^{-4})</td>
<td>(1508)</td>
</tr>
<tr>
<td>10000</td>
<td>0.024</td>
<td>0.005</td>
<td>0.016</td>
<td>0.02</td>
<td>37</td>
<td>304 (^a)</td>
<td>7x10^{-4}</td>
<td>3289</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(483)</td>
<td>(3x10^{-5})</td>
<td>(2070)</td>
</tr>
<tr>
<td></td>
<td>0.0005</td>
<td>0.02</td>
<td>6x10^{-18}</td>
<td></td>
<td></td>
<td>100 (^b)</td>
<td>0.007</td>
<td>10000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(105)</td>
<td>(0.009)</td>
<td>(9524)</td>
</tr>
</tbody>
</table>

The expected time scale of operation of Muller’s ratchet under background selection for Y chromosomes and the neo-Y chromosomes of *Drosophila miranda*. Values calculated with equation (3) are labelled (a), and those with equation (8) in Gessler (1995) corrected for a reduced $N$ are labelled (b). $w/w_i$ is the ratio of the expected mean fitness after 1 million generations, and the initial mean fitness. The values in parenthesis are the expected values in the absence of strong mutations obtained by equation (5) in Chapter 1.
In Table 3.5, I assume this value for the mutation rate per mammalian proto-Y chromosome per generation, and an effective population size one order of magnitude below the one I assume for Drosophila. I show the expected time between clicks of the ratchet and the decline in fitness for different values of the selection coefficients. One can see that, if mildly deleterious mutations are as frequent as is assumed, they can accumulate in the presence of strongly deleterious mutations at significant rates and cause a considerable decline in the mean fitness of the population of Y chromosomes in a short evolutionary time. From Table 3.5, one can see that the effect of background selection in increasing the rate of the ratchet is only of importance in the cases where \( n_{\text{m0}} \gg 1 \).

If, in addition to background selection due to strongly deleterious mutations and Muller's ratchet causing accumulation of mildly deleterious mutations, there is a third class of mutations with effects \( s_y \sim 1/N_e \), one expects these mutations to show an accelerated rate of fixation due to the background selection effects of the more strongly deleterious ones (Charlesworth 1994).

### 3.5.4 The accumulation of transposable elements due to Muller's ratchet

Transposable elements (TEs) are sequences that can insert into new locations of the genome. Because of their property of self-replication, they would invade every site into which they can insert if there were no forces opposing transposition. Excision, deleterious effects of mutations induced by the insertion itself, or selection against chromosomal rearrangements created by ectopic recombination between elements, are among the forces thought to balance transposition (Charlesworth and Langley 1989; Charlesworth et al. 1997; Charlesworth et al. 1992a; Charlesworth et al. 1992b; Charlesworth et al. 1994). These forces prevent TEs from occupying
coding regions, and can keep TEs at low frequencies in the sites to which they insert, even outside the coding regions.

The absence of recombination in Y or neo-Y chromosomes has the automatic consequence of reducing ectopic exchange and reducing selection against rearrangements. This would lead to the accumulation of TEs in these chromosomes, as observed in various organisms (Steinemann and Steinemann 1992; Steinemann and Steinemann 1998; Wichman et al. 1992). In addition, an increased rate of accumulation of these elements due to the ratchet and/or background selection can also be responsible for the numerous copies of TEs present in Y and neo-Y chromosomes (Charlesworth et al. 1994). I illustrate the time scale of this by the following example. Consider the specific case of the neo-Y chromosome of *D. miranda*, whose age is estimated to be about 1 MY (Bachtrog and Charlesworth 2000; Yi 2000). Measures of transposition rate have been described for *D. melanogaster*, in which a mean transposition rate per haploid euchromatic genome per generation is estimated to be about 0.12 (Maside et al. 2000). Assuming that the rate of transposition in *D. miranda* is about the same, a rate of 0.015 for insertion of transposable elements that cause very weak mutational effects on fitness due to insertions (for example, into regulatory regions) is possibly a reasonable value for the neo-Y (as assumed in figure 3.5). Multiplicative fitness effects between TEs is also assumed, since synergistic interactions are expected to be weak in the absence of ectopic exchange (Charlesworth et al. 1997).

Suppose that at time zero a fusion that had occurred between the old Y and an autosomal arm was established, forming a new population of neo-Y chromosomes. In figure 3.5, I show the change in mean fitness over 50,000 generations for a population of $10^5$ chromosomes, subject to mutations with effect $s_i \sim 10^4$, occurring at rate 0.015. In addition, I also consider the effect of the presence of much larger effect mutations (these can be deleterious point mutations, or strong deleterious
mutations caused by TEs when they insert into coding regions). The latter may accumulate by a ratchet, or just cause the background selection effect, depending on the strength of their effect. The population of hypothetical neo-Y chromosomes is initially free of any mutations, so its mean fitness at time 0 has the maximum value 1 (I am ignoring the deleterious mutations that could have got fixed during the fixation of the fusion in the population). As time passes, mutations start to build up and the mean fitness starts to decline. It takes about $1/s_e$ generations for the mean fitness, with respect to the weakly selected mutations, to reach its equilibrium value $\exp(-U)$ (Johnson 1999). During this time, successive losses of the least-loaded classes of the weakly deleterious mutations caused by TEs can be observed. Despite this, the mean fitness does not decline much, because the effect of these mutations is extremely small. In the absence of the large-effect mutations, the mean fitness reaches its equilibrium value, $\exp(-0.015) = 0.985$, (solid line in figure 3.5) after approximately 10,000 generations, and then decays very slowly. However, over 50,000 generations I observe more than 400 clicks of the ratchet for these small-effect mutations.

In the presence of mutations with a selection coefficient of 0.01, the rate of decline in mean fitness after its equilibrium value, 0.951, has been reached, is slightly bigger than in their absence, and a bigger number of clicks is observed over the same period of time (more than 600 clicks). When the effect of the large-effect mutations drops from 0.01 to 0.006-0.005, then not only does the rate at which TEs accumulate increase, but also a drastic decline of mean fitness due to the accumulation of the large mutations is observed. Note that, in these circumstances, TEs are accumulating as a result of the accelerated rate of fixation caused by the reduction in effective population size ($f_{op}$) due to the large-effect mutations, even though these mutations actually experience a very slow ratchet (Charlesworth 1996b). If one assumes that the larger mutations are accumulating at the same rate as in the absence of TEs (i.e, that, due to their small fitness effects, TEs do not interfere
with the dynamics of the much larger mutations) the predicted value of the mean fitness after 50,000 generations is 0.736, very close to that observed in the simulations (0.712, for the case where $s_b = 0.006$). Because over this time the number of TEs has increased from virtually zero to hundreds, and excision was ignored in these simulations, the numbers of TEs will be smaller if excision is considered. But, since the excision rate is thought to be much smaller than the transposition rate (Maside et al. 2000) this effect is likely to be small as we show in figure 3.5 (assuming an excision rate of 1/10 the transposition rate). In addition, as the number of elements builds up on neo-Y chromosomes the transposition rate will increase, since the number of transposition events depends on the number of elements. Therefore the speed of the ratchet will not be constant (as assumed in figure 3.5) but will increase over evolutionary time, making the predictions conservative. One therefore expects a very rapid accumulation of TEs by these processes after a neo-Y chromosome is born.

In conclusion, I have quantified the rate of accumulation of deleterious mutations by Muller’s ratchet considering a model for the distribution of mutation effects that is not only mathematically simple, but also seems to be a type of model that fits the data from induced mutation accumulation experiments. From Table 3.5, it is clear that mutations with small effects on fitness can accumulate (and fix) at high rates in the presence of mutations with larger effects, if they occur frequently in large chromosomes or chromosome regions that lack recombination. From figure 3.5, we conclude that an accumulation of TEs (with very weak deleterious effects) by a ratchet-like mechanism is expected to occur prior to the accumulation of mildly deleterious alleles, although such TEs are assumed to cause a much smaller decline in mean fitness, due to their small effects on fitness.
Figure 3.5. The accumulation of transposable elements in a hypothetical population of neo-Y chromosomes. The solid line represents the decline in mean fitness due to TEs in the absence of larger deleterious mutations. Parameter values are as follows: $N = 10^5$, $U_z = 0.015$, $s_z = 0.0001$ and $U_b = 0.035$. The value of $s_b$ is indicated in the figure.

Muller’s ratchet working in combination with background selection can therefore potentially cause a rapid decline in the mean fitness of a large population of Y or neo-Y chromosomes through the fixation of deleterious alleles, over a biologically reasonable time scale. Of course, the operation of these processes is not incompatible with the operation of other processes: hitchhiking of deleterious mutations due to the fixation of strongly advantageous mutations on the Y chromosome, accelerated rate of fixation of deleterious mutations with $s < 1/N_e$ on the Y due to background selection, a lower rate of adaptation of the Y compared with the X chromosome, due to the operation of background selection on the Y, and an
accelerated rate of fixation of deleterious mutations with $s \sim 1/N_e$ due to the Hill-Robertson effect (Charlesworth and Charlesworth 2000; McVean and Charlesworth 2000; Orr and Kim 1998; Rice 1987).
4 Muller's ratchet and the pattern of variation at a neutral locus

This chapter is based on an article submitted to *Genetics* (Gordo, Navarro and Charlesworth). The results involving the simulations of the structured coalescent were contributed mainly by A. Navarro, who wrote and run most of the structured coalescent simulations.

4.1 Introduction

So far I have been concerned with the question: how many generations on average will it take for an asexual population to lose its present best class of individuals? Although the biological importance of the ratchet can be assessed by the quantification of this time, and the associated decline in mean fitness, the conclusion essentially has been that its extremely high sensitivity to small changes in the parameters, together with our lack of precise knowledge of the values of the relevant parameters of mutation and selection (Keightley and Eyre-Walker 1999), make it hard to draw definitive conclusions about the ratchet's role. For example, if the majority of deleterious mutations are mildly deleterious (with \( s << 1\% \)), then, with our present knowledge of the deleterious mutation rate, I have concluded that Muller's ratchet could potentially be a major process driving the degeneration of \( Y \) chromosomes, even in very large populations such as those of Drosophila. But if this is not the case, its operation may be biologically negligible in that context.

One signature of the operation of Muller's ratchet is the fixation of deleterious alleles as a consequence of the recurrent loss of the best class (Bergstrom and Pritchard 1998; Charlesworth and Charlesworth 1997; Higgs and Woodcock 1995). In this chapter, I try to evaluate another signature of its operation. As discussed in the introduction, the elimination of strongly deleterious mutations can substantially reduce variation levels at linked neutral sites - the effect known as "background selection" (Charlesworth et al. 1993a). Background selection, as
classically stated, assumes that no irreversible accumulation of deleterious mutations occurs, which is not the case when the ratchet is turning. I have therefore asked the following question: what is the level and pattern of neutral variation in a population where Muller's ratchet is operating? Observations on neutral variability in asexual populations or on Y chromosomes may detect the signature of processes such as the ratchet (Charlesworth and Charlesworth 2000), so that it is important to have theoretical predictions of what to expect. This question is examined using Monte Carlo stochastic simulations of a neutral locus embedded in a set of selected loci that accumulate deleterious mutations by the ratchet mechanism. Variability at the neutral locus is measured and compared with both analytical and simulated results based on the structured coalescent. Tajima's $D$ statistic (Tajima 1989), commonly used to test deviations from the standard neutral model, is calculated, and its power to reject neutrality in a population undergoing a ratchet mechanism is assessed.

A model in which the population is subject to deleterious mutations with two major types of effects, like the one studied in chapter 3, is also studied and the level of neutral variability under such a model is compared to an analytical approximation.

### 4.2 Simulation Methods

#### 4.2.1. Multi-locus Monte Carlo simulations

Following the simulation work of the chapters 2 and 3, a haploid, non-recombining population of $N$ chromosomes was simulated with the following life cycle: mutation, reproduction and selection. In each generation, mutations to deleterious alleles occur according to a Poisson distribution with mean $U$. Multiplicative fitness effects of the deleterious mutations are assumed. Two kinds of model of the effects of deleterious mutations are considered: a model in which all mutations have the same selection coefficient ($s$), and a model where mutations can
have two types of deleterious effect \( (s_a \text{ and } s_b) \), with \( s_a < s_b \). At a neutral locus, mutations are generated according to the infinite sites model. The neutral locus has 250 neutral sites and new mutations are not allowed to occur at any site that already carries a mutation. The mutation rate is \( \mu \) for the neutral locus. Every \( N \) generations, a sample of size \( n = 25 \) chromosomes is taken from the population, and variability measures at the neutral locus are computed. For each set of parameter values, 10 simulation runs were done, with 5 samples taken per run (no significant correlation between samples was observed). A total of 50 samples was used for the computations.

### 4.2.2 Coalescent simulations

A model in which all mutations have identical selection coefficients \( (s) \) was simulated using the coalescent process. The method is based on the structured coalescent described by Charlesworth et al. (1995), with a slight modification. The program follows the genealogies of a sample of neutral alleles backwards in time. These can move between the different classes of individuals defined by the number of deleterious mutations harbored by each individual. The frequencies of these classes after selection \( (f_i') \), are given by their deterministic expectation, i.e., according to a Poisson distribution with parameter \( \lambda = U (1-s) / s \) when \( n_o > 1 \). When \( n_o = N e^{-\lambda} < 1 \), the distribution is replaced by a shifted Poisson distribution with parameter \( \lambda = U (1-s) / s - k \), where \( k = \min \{i, \ Nf_i' > 1\} \) (Gessler 1995; see below for details). The first step in the program is to generate a transition matrix, \( Q_{ij} \), of the probabilities that an individual with \( i \) mutations in a given generation has an ancestor with \( j \) mutations in the previous generation, according to Equation (3) of Charlesworth et al. (1995), see also Appendix 4A). Then, the haplotypes in which the neutral alleles in the sample are embedded are generated randomly from the
equilibrium distribution after selection (this is different from Charlesworth et al. (1995), who considered the equilibrium distribution after mutation and prior to selection, although this makes no difference for the values of $s$ considered here). After these two preliminary steps, the backward process starts. Every generation, the number of mutations in the ancestor of each individual is obtained randomly by using the probabilities in the matrix $Q_q$ as expected values. When all the ancestors have been assigned to a class, coalescence is allowed to occur between individuals belonging to the same class, with probability $\frac{k_i(k_i-1)}{2}\frac{1}{N_{f_i}}$, where $k_i$ is the number of lineages with $i$ deleterious mutations present in the sample at a given generation, and $N_{f_i}$ is the deterministic equilibrium size of class $i$, after selection. The possibility of more than one coalescent event within a class is neglected. Simultaneous coalescent events are possible in the same generation if they occur in different classes. Once the most recent common ancestor of the whole sample is reached, neutral mutations are distributed over the gene tree generated by the simulation, according to the infinite sites model, and variability measures are calculated following standard coalescent procedures (Hudson 1990). For each set of parameter values, $10^3$ trees were generated.

4.2.3 Measures of genetic diversity at a neutral locus

Two measures of genetic variation in a sample of alleles at the neutral locus are considered: the mean number of pairwise differences between randomly sampled sequences, $k$; and the number of segregating sites, $S$. Under the infinite-sites model in the absence of deleterious mutations, the expectations of these quantities for a haploid population are as follows (Ewens 1979):

$$k_0 = \theta$$
where \( \theta = 2N\mu \) and \( n \) is the sample size. The 0 subscript refers to the strictly neutral model. In the absence of recombination, these expectations are reduced approximately by a factor \( f_0 = \exp(-UL/s) \) in a large population that is at equilibrium between recurrent mutation to strongly deleterious alleles and their elimination by purifying selection (background selection). This approximation was previously shown to be accurate in a population where Muller’s ratchet does not operate (Charlesworth et al. 1993a).

### 4.2.4. The mutational frequency spectrum

Selection against deleterious mutations is expected to affect \( k \) more than \( S \), since \( k \) is weighted towards variants at intermediate frequencies (Charlesworth et al. 1993a; Tajima 1989). This is observed when computing statistics, such as Tajima’s \( D \), designed to test deviations from the frequency spectrum predicted under strict neutrality. Tajima’s \( D \) is defined as:

\[
D = \frac{k - \theta_w}{\sqrt{\text{Var}(k - \theta_w)}}
\]

where \( \theta_w = S \sum_{i=1}^{n-1} \frac{1}{i} \) and \( \text{Var}(k - \theta_w) \) is calculated by assuming no recombination (Tajima 1989). Negative values of \( D \) are associated with a skew in the distribution of frequencies of neutral mutations towards an excess of rare variants. Because a negative \( D \) is expected in the presence of purifying selection (Charlesworth et al. 1993a), one asks how often, in the presence of a ratchet, can one reject the neutral model because of very negative \( D \)s. For a given \( \theta \) value, I ran standard coalescent
simulations (Hudson 1990) of the neutral infinite sites mutational model and calculated the critical values (at the 95% CI) of the statistic $D$. I then used forward simulations to compute the rejection power, given by the proportion of forward simulations whose observed $D$ was lower than the critical value expected under neutrality: $Pow_1$. Although in some species, such as $Drosophila melanogaster$, one has information about genome wide levels of variability, from which one can estimate $\theta$, in others such information is not available. Because $\theta$ is generally not known in these cases, a power analysis assuming a fixed number of segregating sites, $S$ (Hudson 1993) was also performed. Standard neutral genealogies were generated, $S$ mutations distributed onto them, and the 95% critical values of $D$ were obtained. Afterwards, structured coalescent simulations were run, also distributing $S$ mutations onto the genealogical trees, and $Pow_2$ was calculated. $Pow_2$ is the proportion of times the value of $D$ obtained in the structured coalescent simulations was lower than the critical $D$ in the neutral simulations. $10^4$ genealogical trees were run for every set of parameters.

4.3. Results

4.3.1. Muller’s ratchet and genetic diversity

Suppose that the accumulation of deleterious alleles is occurring due to the repetitive “clicks” of Muller’s ratchet. What is the expected level of variability at a locus evolving neutrally? In figure 4.1, I show the reduction in the mean number of pairwise differences caused by deleterious mutations, i.e., the ratio of the observed $k$ to that expected in the absence of purifying selection, $k_0$, as a function of $s$. I also plot the deterministic equilibrium frequency of the least loaded class, $f_0$. For any value of $N$, with a sufficiently large value of $s$ the reduction in $k$ is independent of $N$ and it is very well approximated by $f_0$. With recurrent mutations with very large
deleterious effects, the rate at which the ratchet operates is extremely low (if it operates at all), and the level of variation at a neutral locus reflects the size of the class of individuals with the highest fitness. This is because any variant arising in less fit classes is quickly driven to extinction (Fisher 1930, pg. 122), or, putting it in another way, any neutral variant sampled from mutated classes is very recently derived from gametes belonging to the fittest class. Thus, for these cases one recovers the classical background selection approximation $E(k) \approx 2f_0N\mu$ (Charlesworth et al. 1993a).

**Figure 4.1.** Simulations of the effect of Muller's ratchet on the mean number of pairwise differences $k$ (relative to that under strict neutrality, $k_0$) as a function of the selection coefficient against deleterious mutations ($s$). The mutation rate is 0.05, $N$ is 1000 (triangles), 3000 (circles) and 8000 (squares). We start observing clicks of the ratchet when $s < 0.03$ for $N = 1000$, $s < 0.02$ for $N = 3000$ and $s < 0.015$ for $N = 8000$. The continuous line is the frequency of the least-loaded class at the mutation-selection deterministic equilibrium, $f_0$. The error bars correspond to two standard errors.
For intermediate selection coefficients, Muller's ratchet starts to operate at a reasonable rate. Two phenomena start to occur: the size of the best class fluctuates around its deterministic equilibrium value and is driven to 0 with a time-lag that varies stochastically (Haigh 1978), and fixations of deleterious alleles in the population start to occur (Charlesworth and Charlesworth 1997; Higgs and Woodcock 1995). Therefore, neither the size of the least-loaded class is constant nor the frequency of every deleterious allele is predicted by the mutation-selection deterministic equilibrium. For example, for $s = 0.01$, and $U = 0.05$, as in figure 4.1, the ratchet clicks on average every 110 generations for $N = 1000$, every 170 generations for $N = 3000$ and every 331 generations for $N = 8000$. One finds that, under conditions favorable to the operation of Muller's ratchet, variability is always higher than the value predicted by $f_0$, and that the reduction in variability is dependent on the value of $N$.

The conditions for the operation of the ratchet require that $s$ is not very large and/or $Nf_0$ is relatively small. This implies that the mean time that a gamete with a deleterious mutation persists in the population can be larger than the mean coalescent time within the least loaded class ($1/s \gg Nf_0$), which means that more loaded classes can significantly contribute to variability. Hence, in these circumstances the relative genetic diversity ($k / k_0$) is higher than the value predicted simply by $f_0$, as seen in figure 4.1. For a given $N$ and $U$, there is a value of $s$, $s_{\text{min}}$, that produces a minimum in diversity. If I plot the results of figure 4.1 as a function of $Nf_0s$, I observe that the minimum occurs for $Nf_0s$ around 1. When $Nf_0s \gg 1$, increasing $s$ increases diversity through the increase in $f_0$; when $Nf_0s \ll 1$, decreasing $s$ increases diversity due to the contribution of classes other than the least-loaded one.

With very weak selection, the reduction in variability becomes very small, and negligible in the limiting case $s \ll 1 / N$, as deleterious alleles then become
effectively neutral and do not interfere with the dynamics of the linked neutral locus at which variation is being measured (Crow and Kimura 1970).

I have tried to approximate the reduction in genetic diversity as follows. Because \(E(k) = 2\mu T_2\), where \(T_2\) is the expected time to the most recent common ancestor of two randomly sampled gametes, I approximate \(T_2\) using the coalescent approach of Hudson and Kaplan (1994) by assuming that a population subject to recurrent deleterious mutations can be thought of as a subdivided population in which mutation plays the role of migration. Under conditions where an approximate mutation-selection balance can be attained, i.e., when \(n_0 >> 1\), (Gessler 1995; Stephan et al. 1993), the sizes of the mutational classes are close to their deterministic expectation most of the time, and the assumption of mutation—selection balance to calculate the coalescent time should produce reasonable results. The expressions for the mean coalescent time are given in Appendix 4A, which are equivalent to Equation (12) of Hudson and Kaplan (1994).

When \(Nf_0 < 1\), the distribution of mutations can deviate considerably from a Poisson with mean \(U/s\). For these cases, Gessler (1995) has suggested that a shifted Poisson distribution of mean \(\lambda\), where \(\lambda = U/s - k\) with \(k = \min \{i, Nf_i > 1\}\), is a better approximation. For these cases I replace Equation (A2) of Appendix A by

\[
Q_{i,i-1} = \frac{U}{U + (1-s) \frac{\lambda}{i}}
\]  

where \(Q_{i,i-1}\) is the probability that a gamete with \(i\) mutations derives from a gamete with \(i-1\) mutations. \(T_2\) can be calculated in the same way as before, but using Equation (1).

In figure 4.2, I compare the results of these analytical approximations (leading to Equation (A5) in the Appendix) with those from the exact Monte Carlo
forward simulations. The deleterious mutation rate is 0.05, and two values of $s$ are considered: $s = 0.005$ and $s = 0.015$. Figure 4.2 shows that variability is more reduced for larger values of $N$ and slowly approaches the value $f_0$ as $N \to \infty$ (absence of the ratchet). The analytical expressions provide reasonably good approximations to the simulation results. Note that, for the case $s = 0.005$ in the range of values of $N$ considered, the deterministic value of $n_0$ is $< 1$, so that Equation (1) was used to calculate the mean coalescent time.

![Figure 4.2](image)

**Figure 4.2.** Relation between the mean number of pairwise differences, relative to that under strict neutrality, and population size ($N$). The deleterious mutation rate is 0.05. The simulation results for $s = 0.005$ (squares) and 0.015 (circles) are shown. The dashed and full lines are the corresponding theoretical values calculated using Equations (A2) - (A5). For sufficiently large values of $N$, $k/k_0$ would be $\sim 5 \times 10^{-4}$ for $s = 0.005$ and $\sim 0.04$ for $s = 0.015$. The error bars correspond to two standard errors.

Simulations of the structured coalescent were run and compared with both the results of the exact Monte Carlo forward simulations and with Equation (A5). As expected, no difference is observed between the mean number of pairwise
differences predicted by Equation (A5) and the one obtained in the coalescent simulations, since they are based on the same assumptions (results not shown).

In figure 4.3, I consider the effects of different values of \( U \) with a constant population size of 2,000 individuals. As expected, the larger the value of \( U \), the bigger the reduction in variability, for any given value of \( s \). The reduction in expected variability predicted by the coalescent approach is a reasonably good approximation to the means obtained in the forward simulations, even for cases where \( n_0 < 1 \). However, for the cases where \( n_0 \ll 1 \), with the smaller values of \( s \) and large values of \( U \) in figure 4.3, coalescent predictions (and, therefore, coalescent simulations) underestimate the mean pairwise differences in the forward simulations.

![Figure 4.3](image_url)

**Figure 4.3.** The dependence of the relative reduction in mean number of pairwise differences, \( k/k_0 \), on the deleterious mutation rate and selection coefficient, with \( N = 2000 \). Squares are the simulation results for \( U = 0.01 \) and the full line is the analytical prediction; circles and the dashed-dot line are the simulation and analytical results for \( U = 0.05 \); diamonds and the dashed line are the results for \( U = 0.1 \). Error bars represent two standard errors. Clicks of the ratchet were observed in the simulations when \( s < 0.01 \) for \( U = 0.01 \), \( s < 0.025 \) for \( U = 0.05 \) and \( s < 0.04 \) for \( U = 0.1 \).
For example, in figure 4.3 with $U = 0.1$ and $s = 0.003$, the reduction in the mean number of pairwise differences observed in forward simulations is 0.217 (with 95% CI 0.027) while the prediction from Equation (A5) is 0.164. A similar behaviour is detected upon close examination of figure 4.2, although the difference there is much smaller. There are at least two reasons to expect a discrepancy between the coalescence approximations and the forward simulations in these cases. The first is that, when $n_0 \ll 1$, the time between clicks of the ratchet is so small that it is very difficult to maintain the stability assumed in the approximations over reasonable periods of time. The second is that, due to this fact, the frequency of the least-loaded class experiences large fluctuations, and spends a considerable amount of time above the expected value assumed in the coalescent approximations. This implies that the level of genetic diversity is likely to be underestimated by the coalescent approach. One does observe such underestimation whenever selection is very weak and the mutation rate is very high, so that the ratchet clicks more than 100 times over $N$ generations.

From the results presented here, I conclude that Muller’s ratchet can considerably reduce genetic diversity at a neutral locus. The extent to which this variation is reduced depends strongly on $s$ (figure 4.1) on $N$ (figure 4.2) and on $U$ (figure 4.3). Interestingly, for large values of $U$ the reduction is essentially unaffected by changes in $s$ over a wide range of intermediate selection coefficients (figure 4.3).

### 4.3.2. Muller's ratchet and the frequency spectrum

I now consider the effect of Muller’s ratchet on the frequency spectrum of mutations at the neutral locus. As explained above, this is examined by calculating Tajima’s $D$ statistic, the test statistic most widely used for this purpose (Fu 1997). In Table 4.1, I show, for different values of $s$, the time between clicks of the ratchet and
Table 4.1
Mean Tajima’s D and power with fixed $\theta$

<table>
<thead>
<tr>
<th>$s$</th>
<th>$T$ (gen)</th>
<th>$D_{coal}$ (2SE)</th>
<th>$D_{forw}$ (2SE)</th>
<th>Pow1 (%)</th>
<th>$T$ (gen)</th>
<th>$D_{coal}$ (2SE)</th>
<th>$D_{forw}$ (2SE)</th>
<th>Pow1 (%)</th>
<th>$T$ (gen)</th>
<th>$D_{coal}$ (2SE)</th>
<th>$D_{forw}$ (2SE)</th>
<th>Pow1 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.003</td>
<td>50</td>
<td>-0.97</td>
<td>-1.01</td>
<td>25</td>
<td>56</td>
<td>-0.96</td>
<td>-1.05</td>
<td>24</td>
<td>73</td>
<td>-1.02</td>
<td>-1.03</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.18)</td>
<td></td>
<td></td>
<td>(0.04)</td>
<td>(0.17)</td>
<td></td>
<td></td>
<td>(0.04)</td>
<td>(0.23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.005</td>
<td>72</td>
<td>-1.08</td>
<td>-1.26</td>
<td>30</td>
<td>80</td>
<td>-1.11</td>
<td>-1.02</td>
<td>31</td>
<td>102</td>
<td>-1.10</td>
<td>-1.05</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.15)</td>
<td></td>
<td></td>
<td>(0.04)</td>
<td>(0.19)</td>
<td></td>
<td></td>
<td>(0.04)</td>
<td>(0.20)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>144</td>
<td>-1.20</td>
<td>-1.30</td>
<td>35</td>
<td>165</td>
<td>-1.21</td>
<td>-1.14</td>
<td>35</td>
<td>293</td>
<td>-1.14</td>
<td>-1.03</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.16)</td>
<td></td>
<td></td>
<td>(0.04)</td>
<td>(0.13)</td>
<td></td>
<td></td>
<td>(0.04)</td>
<td>(0.17)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.015</td>
<td>325</td>
<td>-1.22</td>
<td>-0.96</td>
<td>33</td>
<td>611</td>
<td>-1.15</td>
<td>-1.11</td>
<td>30</td>
<td>2239</td>
<td>-0.90</td>
<td>-0.80</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.22)</td>
<td></td>
<td></td>
<td>(0.04)</td>
<td>(0.18)</td>
<td></td>
<td></td>
<td>(0.06)</td>
<td>(0.18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.02</td>
<td>1292</td>
<td>-1.03</td>
<td>-1.15</td>
<td>27</td>
<td>3478</td>
<td>-0.91</td>
<td>-0.95</td>
<td>23</td>
<td>-</td>
<td>-0.55</td>
<td>-0.65</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.20)</td>
<td></td>
<td></td>
<td>(0.06)</td>
<td>(0.28)</td>
<td></td>
<td></td>
<td>(0.06)</td>
<td>(0.23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.03</td>
<td>-</td>
<td>-0.60</td>
<td>-0.74</td>
<td>13</td>
<td>-</td>
<td>-0.53</td>
<td>-0.80</td>
<td>12</td>
<td>-</td>
<td>-0.53</td>
<td>-0.80</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.20)</td>
<td></td>
<td></td>
<td>(0.06)</td>
<td>(0.22)</td>
<td></td>
<td></td>
<td>(0.06)</td>
<td>(0.23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.04</td>
<td>-</td>
<td>-0.33</td>
<td>-0.29</td>
<td>9</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The mean values of Tajima’s $D$ for different $s$ values (two standard errors are shown below the means), based on forward Monte Carlo simulations ($D_{forw}$) and coalescent simulations ($D_{coal}$). Pow1 is the percentage of coalescent simulations that yielded values of $D$ lower than the critical value (at the 95% probability level) obtained as explained in Methods. $T$ is the time between turns of Muller’s ratchet calculated from the simulations. The symbol - means that no clicks were observed in the simulations. Sample size is 25. $\theta = 2N\mu = 8$, for the neutral locus as a whole. The deleterious mutation rate is 0.05.
Table 4.2
Mean Tajima’s $D$ and power with fixed $S$

<table>
<thead>
<tr>
<th>$s$</th>
<th>$N = 2000$</th>
<th>$N = 3000$</th>
<th>$N = 8000$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$D_{coal}$ (2SE)</td>
<td>Pow2 (%)</td>
<td>$D_{coal}$ (2SE)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$S = 30$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.003</td>
<td>-1.17 (0.01)</td>
<td>33 (0.01)</td>
<td>-1.31 (0.01)</td>
</tr>
<tr>
<td>0.005</td>
<td>-1.33 (0.01)</td>
<td>41 (0.01)</td>
<td>-1.43 (0.01)</td>
</tr>
<tr>
<td>0.01</td>
<td>-1.53 (0.01)</td>
<td>57 (0.01)</td>
<td>-1.66 (0.01)</td>
</tr>
<tr>
<td>0.015</td>
<td>-1.60 (0.01)</td>
<td>65 (0.01)</td>
<td>-1.61 (0.01)</td>
</tr>
<tr>
<td>0.02</td>
<td>-1.31 (0.01)</td>
<td>42 (0.01)</td>
<td>-1.20 (0.01)</td>
</tr>
<tr>
<td>0.03</td>
<td>-0.71 (0.02)</td>
<td>15 (0.02)</td>
<td>-0.59 (0.02)</td>
</tr>
<tr>
<td>0.04</td>
<td>-0.43 (0.02)</td>
<td>9 (0.02)</td>
<td>-0.36 (0.02)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$S = 10$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.003</td>
<td>-1.03 (0.01)</td>
<td>27 (0.01)</td>
<td>-1.17 (0.01)</td>
</tr>
<tr>
<td>0.005</td>
<td>-1.19 (0.01)</td>
<td>35 (0.01)</td>
<td>-1.26 (0.01)</td>
</tr>
<tr>
<td>0.01</td>
<td>-1.36 (0.01)</td>
<td>46 (0.01)</td>
<td>-1.47 (0.01)</td>
</tr>
<tr>
<td>0.015</td>
<td>-1.40 (0.01)</td>
<td>48 (0.01)</td>
<td>-1.43 (0.01)</td>
</tr>
<tr>
<td>0.02</td>
<td>-1.17 (0.02)</td>
<td>34 (0.02)</td>
<td>-1.06 (0.02)</td>
</tr>
<tr>
<td>0.03</td>
<td>-0.65 (0.02)</td>
<td>15 (0.02)</td>
<td>-0.53 (0.02)</td>
</tr>
<tr>
<td>0.04</td>
<td>-0.37 (0.02)</td>
<td>9 (0.02)</td>
<td>-0.31 (0.02)</td>
</tr>
</tbody>
</table>

The mean values of Tajima’s $D$ for different $s$ values (two standard errors are shown below the means), based on simulations of the structured coalescent. Pow2 is the percentage of simulations that yielded values of $D$ lower than the critical value (at the 95% probability level) obtained by coalescent simulations of the neutral model, as explained in Methods. The number of segregating sites, $S$, in the sample is fixed. $U = 0.05$. 

89
average values of $D$ obtained from forward and coalescent simulations, assuming a fixed value of $\theta$. Coalescent simulations were run to compare the results obtained from forward simulations.

As can be seen in Table 4.1, the agreement between forward and coalescence simulations is quite good. The power to detect deviations from neutrality in samples of size 25 is also shown in Table 4.1. Table 4.2 shows the results of coalescent simulations where a fixed number of segregating mutations, $S$, was distributed over the trees (see Methods).

I find that the operation of Muller's ratchet produces negative values of Tajima's $D$ in samples of realistic size. The mean value of $D$ for different values of $N$ and intermediate values of $s$ is around $-1$. For the $\theta$ and $S$ values considered, with intermediate values of $Ns$ there is considerable power to detect deviations from neutrality in samples of size 25. For sample sizes of 10, however, I generally found no power to reject neutrality (results not shown). For a given $N$, the maximum negative average value of Tajima's $D$ is observed for intermediate values of $s$. I observe that, as the time between turns of the ratchet becomes very large, by increasing $s$ and $N$ (or decreasing $U$), the average value of Tajima's $D$ becomes less negative and the frequency spectrum becomes closer to that expected under neutrality ($D \to 0$), as expected from previous results on background selection Charlesworth et al. (1995).

4.4. Muller's ratchet with two types of deleterious mutations

Assume now that there are two major types of deleterious mutations: one class of mutations causing very strongly deleterious effects ($s_b$), and another class with weak deleterious effects ($s_e$), occurring at rates $U_b$ and $U_s$, respectively. Although this mutational model is probably too simplistic biologically, it has been
suggested that it provides a reasonably good fit to data from experiments on the fitness effects of induced mutations, at least in *C. elegans* (Davies et al. 1999). In addition, it allows us to explore the combined operation of two processes: Muller’s ratchet and background selection (see below) (Charlesworth 1996b). As discussed in chapter 3, the deterministic equilibrium frequency of the class with *i* mutations of effect *s*, and *j* mutations of effect *s*₂, after selection, is the product of the relevant Poisson distributions (Johnson 1999; Johnson 2000). In particular the size of the least-loaded class, after selection, is:

\[ n'_{00} = f^*_0 f^*_0 = \frac{U_s(1-s_i)}{s_i} e^{s_2} \]

One can easily extend the coalescent approach used above to this two type of mutations model. The expression for the mean number of pairwise differences relative to the neutral case is given in Appendix 4B.

If \( n'_{00} > 1 \), the population will be close to the deterministic equilibrium most of the time and the sizes of the classes can be well approximated by Equation (B3). When *s* is small and/or \( U_s \) is large, such that \( n'_{00} < 1 \), I approximate the distribution of the classes with respect to these mutations by a shifted Poisson with parameter \( \lambda_p \) (see Appendix 4B) (Gessler 1995).

In Table 4.3, I show the mean number of pairwise differences relative to the neutral case, in populations of size 3,000 and 6,000 subject to both types of deleterious mutations. I also show the case when the deleterious mutations with selection coefficient *s*₂ are absent, for comparison, and the results from Equation (B3), which are referred to as “Theoretical”. The distortion of the neutral frequency spectrum, as measured by the mean Tajima’s *D*, is given for every set of parameters.
Table 4.3

The reduction in the mean number of pairwise differences \( (k/k_0) \) due to Muller’s ratchet with two classes of mutations.

<table>
<thead>
<tr>
<th>( s_x )</th>
<th>( k/k_0 ) ( (2SE) )</th>
<th>( k/k_0 ) Theoretical</th>
<th>( T_{S_x} )</th>
<th>( T_{S_b} ) ( (2SE) )</th>
<th>( D ) ( (2SE) )</th>
<th>( Pow1 ) %</th>
<th>( k/k_0 ) ( (2SE) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>a 0.003</td>
<td>0.124 ( (0.023) )</td>
<td>0.136 ( (0.21) )</td>
<td>44</td>
<td>-</td>
<td>-0.91 ( (0.17) )</td>
<td>16</td>
<td>0.156 ( (0.019) )</td>
</tr>
<tr>
<td>b 0.005</td>
<td>0.116 ( (0.019) )</td>
<td>0.110 ( (0.17) )</td>
<td>57</td>
<td>-</td>
<td>-0.95 ( (0.17) )</td>
<td>22</td>
<td>0.140 ( (0.020) )</td>
</tr>
<tr>
<td>c 0.01</td>
<td>0.082 ( (0.018) )</td>
<td>0.080 ( (0.21) )</td>
<td>115</td>
<td>6753</td>
<td>-0.93 ( (0.19) )</td>
<td>25</td>
<td>0.104 ( (0.024) )</td>
</tr>
<tr>
<td>d 0.015</td>
<td>0.086 ( (0.016) )</td>
<td>0.064 ( (0.24) )</td>
<td>231</td>
<td>3858</td>
<td>-0.98 ( (0.24) )</td>
<td>36</td>
<td>0.102 ( (0.017) )</td>
</tr>
<tr>
<td>e 0.02</td>
<td>0.083 ( (0.018) )</td>
<td>0.067 ( (0.24) )</td>
<td>494</td>
<td>4260</td>
<td>-0.77 ( (0.24) )</td>
<td>21</td>
<td>0.125 ( (0.018) )</td>
</tr>
<tr>
<td>f 0.03</td>
<td>0.085 ( (0.022) )</td>
<td>0.093 ( (0.19) )</td>
<td>4560</td>
<td>4913</td>
<td>-0.93 ( (0.19) )</td>
<td>25</td>
<td>0.193 ( (0.022) )</td>
</tr>
</tbody>
</table>

| g 0.003 | 0.195 \( (0.035) \) | 0.211 \( (0.23) \) | 402 | - | -0.65 \( (0.23) \) | 16 | 0.273 \( (0.036) \) |
| h 0.005 | 0.164 \( (0.024) \) | 0.158 \( (0.23) \) | 656 | - | -0.86 \( (0.23) \) | 24 | 0.291 \( (0.045) \) |
| i 0.01 | 0.15 \( (0.032) \) | 0.172 \( (0.24) \) | 4313 | - | -0.86 \( (0.24) \) | 30 | 0.401 \( (0.056) \) |
| j 0.02 | 0.246 \( (0.039) \) | 0.233 \( (0.28) \) | - | - | -0.36 \( (0.28) \) | 14 | 0.565 \( (0.089) \) |

| k 0.005 | 0.079 \( (0.018) \) | 0.059 \( (0.23) \) | 69 | - | -0.86 \( (0.23) \) | 26 | 0.081 \( (0.011) \) |
| l 0.01 | 0.060 \( (0.017) \) | 0.044 \( (0.21) \) | 154 | 11657 | -0.79 \( (0.21) \) | 16 | 0.051 \( (0.013) \) |
| m 0.02 | 0.062 \( (0.014) \) | 0.049 \( (0.190 \) | 1508 | 15188 | -0.99 \( (0.190 \) | 26 | 0.090 \( (0.019) \) |
| n 0.03 | 0.056 \( (0.013) \) | 0.081 \( (0.25) \) | - | - | -0.65 \( (0.25) \) | 19 | 0.171 \( (0.030) \) |

| o 0.003 | 0.135 \( (0.021) \) | 0.156 \( (0.16) \) | 68 | 480 | -1.18 \( (0.16) \) | 32 | 0.190 \( (0.030) \) |
| p 0.005 | 0.111 \( (0.024) \) | 0.123 \( (0.20) \) | 101 | 325 | -1.14 \( (0.20) \) | 30 | 0.186 \( (0.030) \) |

The symbol - means that no clicks were observed during the runs. \( T_{S_b} \) and \( T_{S_x} \) are the average time (in generations) between clicks of the ratchet with respect to each type of deleterious mutations. Other symbols are as in Table 4.1. \( Pow1 \) is based on forward simulations with fixed \( U = 6 \). The last column contains the reduction of neutral diversity in the absence mutations with selection coefficient \( s_b \).
There are several distinct cases that can occur in a two-type mutational model. The first of such cases is the accumulation of mutations of effect $s$, in the presence of much more strongly deleterious mutations, for which there is no ratchet – i.e. the combined operation of Muller’s ratchet and background selection (as studied in chapter 3). The large effect mutations are expected to reduce variability by a fraction $f_{ob}$, and the additional presence of the other mutations, which are accumulating due to Muller’s ratchet, is expected to reduce variability even more.

For the cases where this occurs (a, b, g, h, i, k in Table 4.3), I observe that Equation (B3) gives good predictions of the relative diversity observed in the simulations. Strongly deleterious mutations reduce diversity at neutral sites by a fraction $f_{ob}$, but, as has been seen in the previous chapter, they also reduce the effective population size experienced by the small effect mutations by approximately the same amount. The small effect mutations will then cause a reduction in genetic diversity according to this new effective size $\left( Nf_{ob} \right)$. It follows that, in this case, the resulting reduction in the mean number of pairwise differences caused by both types of mutations is given by:

$$ f_{ob} \frac{k}{k_0} \left( Nf_{ob}, U, s \right) $$

with $k / k_0$ calculated with Equation (A5).

The average values of Tajima’s $D$ are around $-0.9$ and there is some power to reject neutrality in samples of reasonable size (25 chromosomes and $\theta = 6$, in the cases in Table 4.3).

The second case occurs when both type of mutations are accumulating due to the ratchet. In Table 4.3, I show some examples of this (c,d,e,f,l,m,o,p). I observe that Equation (B3) predicts the expected mean number of pairwise differences relative to that under strict neutrality reasonably well. Average values of Tajima’s $D$
are between −0.8 and −1, for the θ value considered, and there is some power to detect a distortion in the frequency spectrum, for a sample size of 25.

The third case occurs if the effects of both types of mutations are very large and/or the mutation rates are very small, such that none will accumulate. This corresponds to the classical background selection model, with no recombination, and two mutational classes. In Table 4.3, one observes that, when one does not observe any clicks of the ratchet (cases j and n) and when \( n_{oo} >> 1/s_i \) and \( n_{oo} >> 1/s_b \), the reduction in genetic diversity is well approximated by \( f_{oo} \) (as expected from the expressions in Appendix B). Note that this is the result expected from a one-class deleterious mutational model in which the relevant selection coefficient is the harmonic mean of the selection coefficients in the two-class mutational model (Charlesworth 1996a).

The fourth case occurs when the presence of strongly deleterious mutations reduces the effective population size by such a large amount that the smaller mutations become effectively neutral, i.e. \( N_f s < 1 \) (Charlesworth 1996b). Under these conditions, genetic drift is the major force determining the dynamics of the small effect mutations and driving them to fixation. Some examples of this case are considered in Table 4.4, with two different mutational models for the small-effect mutations: one considering irreversible mutation, and another, more realistic model, allowing for back-mutation (McVean and Charlesworth 2000).

In these cases, the reduction in the mean number of pairwise differences is very close to the one caused by the strong mutations, since the weak ones are effectively neutral (Kimura 1983) and do not have any significant effect on the neutral locus at which variation is being measured. Therefore, \( k / k_0 \) can essentially be approximated by \( f_{ob} \). In this case, the average Tajima’s \( D \) is much less negative than in some of the previous cases, and it is very difficult to detect distortions in the
frequency spectrum (cf. Charlesworth et al. 1995), especially when allowing for back-mutation.

**Table 4.4**
The effects of background selection on weakly selected mutations with and without back-mutation.

<table>
<thead>
<tr>
<th>$U_{back}$</th>
<th>$s_p$</th>
<th>$k/k_0$</th>
<th>$f_{ob}$</th>
<th>$D$</th>
<th>$P_{ow1}$</th>
<th>$k/k_0$</th>
<th>$D$</th>
<th>$P_{ow1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(2SE)</td>
<td>(2SE)</td>
<td></td>
<td>(%)</td>
<td>(2SE)</td>
<td>(2SE)</td>
<td>(%)</td>
</tr>
<tr>
<td>Without back-mutation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 0.001</td>
<td>0.127</td>
<td>0.115</td>
<td>-0.76</td>
<td>22</td>
<td>0.130</td>
<td>-0.46</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>(0.032)</td>
<td>(0.24)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 0.0005</td>
<td>0.102</td>
<td>0.115</td>
<td>-0.61</td>
<td>21</td>
<td>0.130</td>
<td>-0.46</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>(0.025)</td>
<td>(0.26)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With back-mutation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.001 0.0005</td>
<td>0.115</td>
<td>0.115</td>
<td>-0.46</td>
<td>14</td>
<td>0.130</td>
<td>-0.46</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>(0.027)</td>
<td>(0.29)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01 0.0005</td>
<td>0.146</td>
<td>0.115</td>
<td>-0.39</td>
<td>8</td>
<td>0.130</td>
<td>-0.46</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>(0.029)</td>
<td>(0.22)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$U_{back}$ is the mutation rate to back-mutations, with respect to the small effect mutational type. Power is based on forward simulations with $\theta = 6$. Other symbols as in Table 4.1. Parameter values are: $N = 3000$, $U_b = 0.09$, $s_b = 0.04$, $U_s = 0.01$. The simulation results just due to the presence of mutations with effect $s_b$ at rate $U_b$ (i.e. without weak mutations) are presented in the last column. The value of $U_b$ was chosen in order to make the effect of background selection sufficiently strong so that the weak mutations become effectively neutral in the presence of the strong mutations.

As in the previous model, in this two-type of mutational model one also observes that, when Muller’s ratchet starts to operate, the level of $k/k_0$ is roughly the same across intermediate values of the selection coefficient, for a fixed population size and mutation rate.

**4.5. Discussion**

**4.5.1. Muller’s ratchet and neutral variation**

Genetic diversity at a neutral locus results from the balance between the rate at which variation is generated (mutation pressure) and the rate with which it is lost.
(genetic drift). In a population that is permanently at equilibrium under recurrent mutation to deleterious alleles, in the absence of recombination neutral genetic diversity is expected to be smaller than the strict neutral expectation (Charlesworth et al. 1993a). This results from the fact that a large fraction of individuals in such population are destined to be eliminated relatively quickly, so that its effective size is reduced to the class of individuals that do not carry deleterious mutations, $N_f_0$ (Charlesworth et al. 1993a).

Here, I have quantified the expected genetic diversity when a population is not permanently at equilibrium, but is losing its least-loaded class at a given rate. I have shown that the operation of Muller’s ratchet causes a considerable reduction in genetic diversity. The extent to which such variation is reduced is a function not only of the relevant mutation and selection parameters but also of population size. In particular, in a population where Muller’s ratchet does not operate, or does so at an exceedingly slow rate, which is expected when $N_f_0 \delta_t >> 10$, the effective size is well approximated by $N f_0$. But when the ratchet starts to operate, the effective size is higher than $N f_0$. Although the operation of Muller’s ratchet has been suggested to cause higher values of genetic diversity than is given by the simple $2N f_0 \mu$ approximation (Charlesworth et al. 1993a), this study is the first attempt to formally demonstrate that it does so, and to estimate by how much.

I have shown that the mean coalescent time of two randomly sampled alleles derived from a structured coalescent model with fixed class size (Hudson and Kaplan 1994) is a good predictor of expected genetic diversity when the ratchet is operating. Just as is observed in the full Monte Carlo simulations, the analytical approximation predicts a minimum genetic diversity for an intermediate value of the selection coefficient. Our results are closely related to those of Higgs and Woodcock (1995), who studied the effect of deleterious mutations on genealogies, in very small populations, and showed that the probability of common parentage was maximal for
some intermediate value of the selection coefficient against deleterious alleles. The results from figure 8 of Higgs and Woodcock (1995) can be obtained by the coalescent approximation proposed here, provided that one corrects them for sampling after selection.

These results are also related to those of Tachida (2000), who observed that diversity at neutral sites was minimal for an intermediate strength of selection, although his model is different from the one considered here. In Tachida’s IMC (independent multicodon) model, a gene is composed of a set of completely linked sites. One-third of the sites are neutral and two-thirds are selected, with selection coefficients drawn from a normal distribution. In our model, the selection coefficient is constant, but the qualitative effect on neutral diversity is the same. The simulation results regarding genetic diversity at neutral sites in Table 1 of Tachida (2000) can be obtained by the coalescent approximation, if I substitute $s$ in the approximation by the value corresponding to the mean strength of selection ($\alpha$) considered in his Table 1 ($\alpha - 2Ns$). With the formula (A5), one obtains good estimates of the average genetic diversity at neutral sites observed in his simulations, except when $\alpha < 5$. As an example for $u = 1 \times 10^{-5}$ per site (implying $U=0.002$ for the whole non-recombining region) and $\alpha = 5$ (implying $s = 0.005$) the value of neutral variability observed in Tachida’s simulations is 0.00857 and the value predicted by the coalescent approximation is 0.00836.

In contrast to the classical background selection model with strong selection (Charlesworth et al. 1995; Hudson and Kaplan 1994), if Muller’s ratchet is operating under weak selection, a considerable distortion of the frequency spectrum at the neutral locus, towards an excess of rare variants, is expected in samples of realistic size (as seen in Table 4.1). But such an effect may be difficult to detect when the ratchet causes a very large reduction in variation, as may be the case in large
populations (see below). This signature of the ratchet is quite close to that of selective sweeps, but not as extreme (see below).

Because a model that considers that all deleterious mutations have the same effect on fitness is certainly a simplification, I also studied the pattern of neutral variation under a two-type deleterious mutational model. I considered several distinct cases. In the case where none of the deleterious mutations accumulates, the classical background selection scenario, I recover the expected prediction: the reduction in genetic diversity is well approximated by considering the harmonic mean of the selection coefficients of the two-type model. Tajima's $D$ is negative on average, but distortions of the frequency spectrum are hard to detect (see Table 4.3). In the case where one of the mutational types is sufficiently strongly selected against that it does not accumulate, but the other type of deleterious mutations does accumulate, I essentially observe the effects of the ratchet in a population of reduced size. In the case where both mutational types accumulate, I again observe negative values of Tajima's $D$ and obtain reasonably good predictions of the genetic diversity by the extended coalescent approach.

4.5.2. Muller's ratchet and the Y chromosome

Because Muller's ratchet has been suggested to be involved in shaping the evolution of Y chromosomes (Charlesworth 1978; Rice 1994) and I have argued in the previous chapters that it is theoretically possible for it to be a potentially major process in causing their degeneration, it is interesting to try to quantify the levels of neutral variation expected under its operation. It is of special interest to ask about the diversity levels expected under the ratchet in systems with relatively young Y chromosomes (Charlesworth and Charlesworth 2000; Rice 1996). As mentioned in chapter 3, some examples of these systems, for which there are some variability
measures, are the Y chromosomes of the plant species *Silene latifolia* and *S. dioica* (Filatov et al. 2000) and the neo-Y chromosomes in some Drosophila species (Bachtrog and Charlesworth 2000; McAllister and Charlesworth 1999; Yi 2000; Yi and Charlesworth 2000).

In figures 4.4 and 4.5 I show some expectations for the signatures of Muller’s ratchet: the expected number of fixations of deleterious alleles over a period of 500,000 generations, figure 4.4A (in the case of Drosophila, this corresponds to approximately 0.1 My), the reduction in the mean number of pairwise differences (figure 4.4B), and the average values of Tajima’s $D$ caused by its operation (figure 4.5A and 4.5B). The number of fixations is estimated by calculating the number of clicks of Muller’s ratchet over such a period of time, since there is a one-to-one correspondence between clicks and fixation events (Charlesworth and Charlesworth 1997; Higgs and Woodcock 1995). The parameter values $N$ and $U$ were assigned in the light of the data presently available (Drake et al. 1998; Filatov et al. 2000; Keightley and Eyre-Walker 1999; Yi and Charlesworth 2000), although as already emphasized these values are currently under debate.

As already pointed in chapter 2 and 3, hundreds of fixations of deleterious alleles in these large populations of non-recombining chromosomes are expected as a consequence of the operation of the ratchet, if a considerable fraction of deleterious mutations have intermediate effects on fitness when heterozygous ($h s << 1\%$). The decline in mean population fitness is maximal for intermediate values of $s$. For example if $N = 125,000$ and $U = 0.01$, with $s = 0.001$, the ratchet reduces mean fitness to 60% of its initial value, after 500,000 generations. With smaller $s$, for example $s = 0.0002$, the value is 74%. Very small values of $s$ imply smaller effects on mean fitness, making the process less effective as a cause of degeneration.
Figure 4.4. The signatures of Muller's ratchet in large hypothetical populations of non-recombining Y chromosomes. The parameters are $N = 125,000$ for dashed lines and $N = 500,000$ for filled lines, with $U = 0.01$ for circles and $U = 0.03$ for squares. (A) The expected number of fixations over a period of 500,000 generations. These are based on the expressions for the time between clicks of Muller's ratchet provided in the Chapter 2 for the cases when $N f_0 > 1$ and the results of Gessler (1995) (Equation 8) for the cases when $N f_0 < 1$. (B) Expected mean number of pairwise differences relative to that in the absence of deleterious mutations, calculated using the analytical prediction.
Figure 4.5. Average value of Tajima’s $D$ for a sample size of 12 and for $\theta = 5$ (A) and sample size 12 for $\theta = 100$ (B). Lines and symbols are as in Figure 4.4.
Associated with the fixation of deleterious mutations by the ratchet, a reduction in genetic diversity of around ten to a hundred-fold is expected (as calculated by Equation (B3) and the simulations of the structured coalescent).

In samples of moderate size \( n = 12 \) in figure 4.5, which corresponds to the sample size for which variation has been studied in *D. miranda* (Bachtrog and Charlesworth 2000), average Tajima’s *D* values of around \(-1\) are expected when the selection coefficient is of intermediate value. For larger samples, the average values of Tajima’s *D* become more negative. As an example, with a sample size of 40, with \( N = 125,000, U = 0.01 \) and \( s = 0.1-0.2\% \), I observed an average value of Tajima’s *D* of \(-1.7\) for \( \theta = 50 \) and \(-1.9\) for \( \theta = 100 \). The power to reject neutrality for these two examples was greater than 80% (assuming a fixed \( \theta \)). A large amount of sequence information and large samples are, however, needed to detect this effect. For example, in Drosophila, where normal levels of variability are around 1-3% per nucleotide site (Moriyama and Powell 1996), the above example implies sequencing around 5,000 to 10,000 neutral sites. One can ask if increasing sample size will produce higher power than increasing \( \theta \) by increasing the number of sites sequenced. From simulations of the structured coalescent with \( n\theta \) held constant (Table 4.5), I observe that increasing \( n \) seems to give more power than increasing \( \theta \).

### Table 4.5

<table>
<thead>
<tr>
<th>( s )</th>
<th>( k / k_0 )</th>
<th>( \theta=5 ) ( n=12 ) ( \theta n=60 )</th>
<th>( \theta=5 ) ( n=60 ) ( \theta n=300 )</th>
<th>( \theta=25 ) ( n=12 ) ( \theta n=300 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0002</td>
<td>0.04</td>
<td>3</td>
<td>32</td>
<td>18</td>
</tr>
<tr>
<td>0.002</td>
<td>0.02</td>
<td>2</td>
<td>28</td>
<td>16</td>
</tr>
<tr>
<td>0.004</td>
<td>0.09</td>
<td>3</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>0.005</td>
<td>0.14</td>
<td>4</td>
<td>8</td>
<td>7</td>
</tr>
</tbody>
</table>

Symbols are as in Table 4.1. \( N = 125,000 \) and \( U = 0.01 \). For this parameters with \( s = 0.004 \) and 0.005 the speed of the ratchet is effectively zero and there is essentially no power to reject neutrally.
If Muller's ratchet is operating in these large populations one expects variability levels to be very low. This is expected across a wide range of values of selection coefficients for which the ratchet can operate, since, under this model, the mutation rate is the major determinant of the level of variation expected (figure 4.4B).

The additional presence of much more strongly deleterious mutations, causing background selection, will result not only in an increase in the number of fixations, as has been seen in chapter 3, but also in a bigger reduction in genetic diversity, as expected from the results presented before. In Table 4.6 I show a quantitative example of what to expect under such model in a large population of size 125,000.

Table 4.6

<table>
<thead>
<tr>
<th>$U_s$</th>
<th>$s_s$</th>
<th>$U_b$</th>
<th>$s_b$</th>
<th>$k/k_0$ no strong mutations</th>
<th>$k/k_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.004</td>
<td>0.01</td>
<td>0.01</td>
<td>0.035</td>
<td>0.090</td>
</tr>
<tr>
<td>0.01</td>
<td>0.002</td>
<td>0.01</td>
<td>0.01</td>
<td>0.013</td>
<td>0.020</td>
</tr>
<tr>
<td>0.01</td>
<td>0.001</td>
<td>0.01</td>
<td>0.01</td>
<td>0.019</td>
<td>0.021</td>
</tr>
<tr>
<td>0.01</td>
<td>0.0005</td>
<td>0.01</td>
<td>0.01</td>
<td>0.025</td>
<td>0.028</td>
</tr>
<tr>
<td>0.01</td>
<td>0.0002</td>
<td>0.01</td>
<td>0.01</td>
<td>0.037</td>
<td>0.042</td>
</tr>
<tr>
<td>0.01</td>
<td>0.001</td>
<td>0.01</td>
<td>0.02</td>
<td>0.020</td>
<td>0.028</td>
</tr>
<tr>
<td>0.01</td>
<td>0.0002</td>
<td>0.01</td>
<td>0.02</td>
<td>0.040</td>
<td>0.042</td>
</tr>
</tbody>
</table>

Population size is 125,000. The last column gives the contribution to the reduction in diversity solely due to the small effect mutations.

There is evidence for reduced levels of variability in some Y chromosome systems (Bachtrog and Charlesworth 2000; Charlesworth and Charlesworth 2000; Filatov et al. 2000; McAllister and Charlesworth 1999; Yi 2000; Yi and Charlesworth 2000). For the Y chromosome of the dioecious plant *S. latifolia*, and
the neo-Y chromosome of *D. miranda*, nucleotide variability is 20-30-fold lower than for the X chromosome (Filatov et al. 2000; Bachtrog and Charlesworth, unpublished results). If the simple process I have studied was the sole cause of the observed reduction, the results in figure 4.4B imply that the deleterious mutation rate for such non-recombining chromosomes is unlikely to be higher than 0.01.

The reduction in variability could of course be caused by another process, such as a recent selective sweep, as mentioned in the first chapter. When an advantageous mutation arises and goes to fixation in a non-recombining population, it wipes out linked neutral variation – the hitchhiking effect (Maynard Smith and Haigh 1974). After such a sweep, variation is slowly restored by mutation, with most of the new neutral variants being at low frequency. Selective sweeps therefore cause distortions of the neutral frequency spectrum (Braverman et al. 1995; Simonsen et al. 1995), just as with repetitive clicks of the ratchet. In Table 4.7, the pattern of variability under the ratchet is compared to that under a recent sweep. I assume knowledge of the neutral equilibrium value of $\theta$, in the absence of any of these processes, and study conditions under which genetic diversity is reduced around 20 – 30 fold. As it is clear from Table 4.7, a recent sweep produces generally more negative average Tajima’s $D$ than the ratchet, for a given reduction in diversity. In small samples a sweep is more likely to be detected than the operation of the ratchet, but it is clear that there is a wide range of parameter space in which no unambiguous conclusion can be drawn. Other statistics, such as patterns of linkage disequilibrium, could also be helpful in trying to distinguish between these and other models (Charlesworth and Charlesworth 2000).
Table 4.7
Comparison of the ratchet with the hitchhiking model

<table>
<thead>
<tr>
<th>Ratchet $N = 125,000$ $U = 0.01$</th>
<th>Hitchhiking $N = 125,000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$S$</td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.79</td>
</tr>
<tr>
<td>40</td>
<td>1.66</td>
</tr>
<tr>
<td>100</td>
<td>3.00</td>
</tr>
<tr>
<td>$\theta = 10$</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1.63</td>
</tr>
<tr>
<td>40</td>
<td>3.38</td>
</tr>
<tr>
<td>100</td>
<td>5.92</td>
</tr>
<tr>
<td>$\theta = 20$</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>3.21</td>
</tr>
<tr>
<td>40</td>
<td>6.78</td>
</tr>
<tr>
<td>100</td>
<td>11.98</td>
</tr>
<tr>
<td>$\theta = 10$</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1.63</td>
</tr>
<tr>
<td>40</td>
<td>3.38</td>
</tr>
<tr>
<td>100</td>
<td>5.92</td>
</tr>
<tr>
<td>$\theta = 20$</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>3.21</td>
</tr>
<tr>
<td>40</td>
<td>6.78</td>
</tr>
<tr>
<td>100</td>
<td>11.98</td>
</tr>
</tbody>
</table>

The reduction in genetic diversity was chosen to be of the order of that observed in the *D. miranda* neo-Y chromosome data (Bachtrog and Charlesworth 2000). The $S$ and $D$ values are means over 1,000 coalescent trees for each model. For the ratchet, results are based on the structured coalescent; for hitchhiking, results are based on coalescent simulations (Hudson 1990) from a population that expanded from size 1 to $N$ instantaneously at time $T_{hh}$ in the past. $T_{hh}$ was chosen to produce the same mean level of variability as with the ratchet.
4.5.3. Muller's ratchet on a non-recombining autosomal region

Since Muller's ratchet is expected to cause different effects for non-recombining chromosomes with different numbers of genes, given that both its speed and its effects on variability are highly dependent on the deleterious mutation rate, I now ask what its effects can potentially be on a much smaller chromosome than the one considered previously. The dot (4th) chromosome of D. melanogaster appears to have very little or no crossing over (Ashburner 1989). Let us assume for the sake of argument that there is no recombination. The 4th chromosome constitutes about 1% of the euchromatic genome of D. melanogaster. This means that the rate of accumulation of intermediate and strongly deleterious (s >> 0.5%) mutations is negligible (unless the deleterious mutation rate is much higher than I have been assuming), and that the reduction of genetic diversity due to background selection caused by such mutations is small.

Charlesworth (1996a) argued that one important cause of reduction of neutral variation in small chromosomal regions with very little recombination is selection against transposable elements. Such selection will balance their transposition, resulting in a stable equilibrium of mean element copy number. This selection can come from at least two factors: selection against insertions and selection against the effects of rearrangements caused by ectopic exchange between elements (Charlesworth et al. 1994). If the population of 4th chromosomes is at transposition-selection equilibrium with respect to TEs, then their effect on linked neutral variability is simply $e^{-\sigma_n}$ where $n_{TE}$ is the mean element copy number (Charlesworth 1996a). However if this equilibrium is not attained and if, for example, the mean number of elements is increasing due to the ratchet (Charlesworth et al. 1994; see discussion in chapter 3) the effect will be different. Below I explore the latter possibility.
Some caution as to be taken in applying the results of the haploid model studied to this genetic system. Unlike the Y chromosome, which is only passed through males and can be thought as a haploid non-recombining population, this hypothetical non-recombining 4th chromosome constitutes a non-recombining sexual population. In order to calculate the potential rate of accumulation of TE's due to the ratchet process, one must now consider the effective size of the species, which for D. melanogaster is around $2 \times 10^6$ (Moriyama and Powell 1996), and the selection coefficient against heterozygous insertions, which is thought to be on the order of $10^4$ (Maside et al. 2000). I will consider different values of the selection coefficient against element insertions around this estimate, and calculate the expected reduction in diversity, as well as the predicted rate of accumulation. There is another very different feature, in relation to the consequences of a ratchet, between a Y chromosome and a 4th chromosome. Because in a haploid asexual populations there is a one-to-one correspondence between clicks of the ratchet and fixation events, each click on a Y chromosome implies the fixation of a deleterious allele in the whole population. But in a diploid non-recombining sexual population this correspondence is only observed if the mutations are close to codominance (Charlesworth and Charlesworth 1997). As explained in Charlesworth and Charlesworth (1997), if the dominance coefficient $h$ is high, a diploid population with no recombination but with segregation (with a size $N$, diploid deleterious mutation rate $U$, selection coefficient against homozygous mutations $s$ and against heterozygous mutations $hs$) will have a similar behaviour to an asexual haploid (with population size $2N$, mutation rate $U/2$ and selection coefficient $hs$).

In Table 4.8, I compare the results of simulations of haploid asexual populations with those that were obtained previously for diploid populations. These are from Table 2 of Charlesworth and Charlesworth (1997) and from Tables 1 and 2 of Charlesworth et al. (1993b).
### Table 4.8
Comparing diploid with haploid populations

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Diploid</th>
<th></th>
<th></th>
<th>Haploid</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>U</td>
<td>h</td>
<td>s</td>
<td>No. fixed</td>
<td>Min No.</td>
<td>SE</td>
<td>Min No.</td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>-----------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>0.05</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>77.2</td>
<td>2.6</td>
<td>62.5</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0</td>
<td>0.545</td>
<td>56.9</td>
<td>1.6</td>
<td>47.4</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>1.5</td>
<td>32.1</td>
<td>2.8</td>
<td></td>
<td>27.4</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>0.35</td>
<td>0.1</td>
<td>11.1</td>
<td>11.4</td>
<td>1.0</td>
<td></td>
<td>14.5</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>0.5</td>
<td>0.1</td>
<td>6.6</td>
<td>6.7</td>
<td>0.7</td>
<td></td>
<td>6.3</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>0.6</td>
<td>0.1</td>
<td>4.5</td>
<td>4.6</td>
<td>0.4</td>
<td></td>
<td>3.3</td>
</tr>
<tr>
<td>25</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>7</td>
<td>43.5</td>
<td>3.3</td>
<td></td>
<td>41.1</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>2.5</td>
<td>32.3</td>
<td>2.6</td>
<td></td>
<td>29.4</td>
</tr>
<tr>
<td>100</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>1.6</td>
<td>22.4</td>
<td>2.3</td>
<td></td>
<td>17.4</td>
</tr>
<tr>
<td>50</td>
<td>0.5</td>
<td>0.2</td>
<td>0.1</td>
<td>10.1</td>
<td>222.2</td>
<td>2.9</td>
<td></td>
<td>228.6</td>
</tr>
<tr>
<td>100</td>
<td>0.5</td>
<td>0.2</td>
<td>0.1</td>
<td>5.5</td>
<td>191.3</td>
<td>2.3</td>
<td></td>
<td>196.1</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.9</td>
<td>49.8</td>
<td>2.0</td>
<td></td>
<td>48.0</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>0.35</td>
<td>0.1</td>
<td>10.2</td>
<td>11.8</td>
<td>2.0</td>
<td></td>
<td>14.5</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>36.8</td>
<td>1.6</td>
<td></td>
<td>29.4</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>2.7</td>
<td>9.6</td>
<td>2.3</td>
<td></td>
<td>11.1</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>0.35</td>
<td>0.2</td>
<td>0.5</td>
<td>0.6</td>
<td>0.3</td>
<td></td>
<td>1.2</td>
</tr>
</tbody>
</table>

The number represent averages over 10 and 20 replicate runs for a period of 2000 generations. No. fixed and Min no. corresponds to the average number of mutations fixed in the whole population and the minimum number of mutations carried by an individual, respectively. Diploid population results are from the simulations by Charlesworth and Charlesworth (1997) and Charlesworth et al. (1993b).

Although it is clear that the results for the haploid population can be used to quantitatively estimate the rate of accumulation, they cannot be used to quantitatively predict the rate of fixation. However, the qualitative conclusion drawn by Charlesworth and Charlesworth (1997) is that, if $h$ is much lower than 0.5 one expects accumulation, whose rate can be calculated using the results on Chapter 2 and 3, but does not expect fixation. Drosophila data suggest that $h$ is negatively correlated with $s$ (Crow 1993). If the majority of TE insertions do cause small effects
on fitness and have $h$ close to 0.5, then the rate of accumulation will correspond to the rate of fixation.

**Table 4.9**

Accumulation and reduction in genetic diversity on a hypothetical 4th chromosome

<table>
<thead>
<tr>
<th>$h_{STE}$</th>
<th>Time (generations)</th>
<th>$\frac{dLn w}{dt}$</th>
<th>$k/k_0$</th>
<th>$f_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00001</td>
<td>1890</td>
<td>-5.3 x 10^{-9}</td>
<td>0.036</td>
<td>7.7 x 10^{-53}</td>
</tr>
<tr>
<td>0.00005</td>
<td>20000</td>
<td>-2.5 x 10^{-9}</td>
<td>0.018</td>
<td>3.8 x 10^{-11}</td>
</tr>
<tr>
<td>0.000075</td>
<td>13333</td>
<td>-5.6 x 10^{-9}</td>
<td>0.015</td>
<td>1.1 x 10^{-7}</td>
</tr>
<tr>
<td>0.0001</td>
<td>8727</td>
<td>-1.2 x 10^{-8}</td>
<td>0.013</td>
<td>6.1 x 10^{-6}</td>
</tr>
<tr>
<td>0.00015</td>
<td>8664</td>
<td>-1.7 x 10^{-8}</td>
<td>0.01</td>
<td>3.4 x 10^{-4}</td>
</tr>
<tr>
<td>0.0002</td>
<td>24415</td>
<td>-8.2 x 10^{-9}</td>
<td>0.01</td>
<td>2.5 x 10^{-3}</td>
</tr>
<tr>
<td>0.0003</td>
<td>3.4 x 10^6</td>
<td>-8.8 x 10^{-11}</td>
<td>0.023</td>
<td>1.8 x 10^{-2}</td>
</tr>
<tr>
<td>0.0004</td>
<td>3.8 x 10^{13}</td>
<td>-1.1 x 10^{-17}</td>
<td>0.053</td>
<td>5.0 x 10^{-2}</td>
</tr>
</tbody>
</table>

$N = 2 \times 10^6$ and $U_{TE} = 0.0012$, $f_0$ is the fraction of chromosomes free of TEs in an effectively infinite population at equilibrium. Time is the time in between clicks of Muller’s ratchet calculated using the results in chapter 2 when $Nf_0 > 1$ and those of Gessler (1995) when $Nf_0 < 1$.

In Table 4.9 I consider a hypothetical non-recombining chromosome of the size of the 4th. I assume that this chromosome is subject to insertions of transposable elements and that there is weak selection against each insertion (since I assume no recombination, I assume that there is no selection against the effects cause by ectopic exchange). I assume that the transposition rate to such a chromosome is the transposition rate per haploid euchromatic genome (Maside et al. 2000) times the fraction of the genome represented by the 4th chromosome (assuming that elements insert at random). This implies a transposition rate of about 0.0012 (i.e. 0.12x1%). I thus treat TE insertions as a source of relatively weak deleterious mutations, and take TEs, irrespectively of the family to which they belong, to have the same transposition
rate. This is surely an approximation since it is known that different element families have different transposition rates (Maside et al. 2000), and excision is ignored, but it will give an idea of their potential combined effect.

The very few studies on levels and patterns of variability in the 4th chromosome seem to point towards a reduced level of variability and a larger mean copy number of elements on the 4th chromosome than in chromosomal regions with normal levels of recombination (Charlesworth et al. 1992b; Bartolome, Maside and Charlesworth, unpublished results). Berry et al. (1991) found no polymorphism in a sample of 10 chromosomes of *D. melanogaster*. Jensen et al. (submitted) found levels of genetic diversity at silent sites of around $5 \times 10^4$, much lower than the 1-3% observed in autosomal regions with high recombination rates (a 20-60 fold reduction in variability). These calculations, although made under the restricted assumptions mentioned, show that a reduction on variability of such order could in principle be produced by weak selection against element insertions. But they also point towards a considerable accumulation of TEs if the deleterious effects against insertions are very weak (of the order of $10^{-4}$), although with no significant decline in mean fitness.

The increased number of elements would be accompanied by fixations if $h$ is near 0.5, but if $h$ is much smaller then fixations are not expected. Sufficient data on patterns of polymorphism of TEs on the 4th chromosome is not yet available. But from table 4.9, I observe that in order to produce a strong reduction in variability, a large number of elements should have accumulated. This appears to be inconsistent with the results presented in Table 4 of Charlesworth et al. (1992b), where, although there is a significant excess of TEs on the 4th chromosome it is not of the magnitude expected under the results on table 4.9.

In addition, it is not completely clear if all of the 4th chromosome constitutes a non-recombining unit, which makes these calculations rather tentative. If there is low but a non-zero rate of recombination, then the ratchet is supposed to be slowed down
and the level of variability expected to increase, and the model studied here will not apply.

Finally, I should mention that it is not yet completely clear what are the major forces controlling element copy number in populations. If, as I have mentioned in chapter 3, ectopic exchange is a much more important force than selection against element insertions in controlling element number, then an increased number of elements is obviously expected in regions of low recombination simply due to this fact.

APPENDIX 4.A

Assume that there are $m$ mutational classes in the population, so that a sample may contain gametes with 0, 1, 2 up to $m$ mutations. The probability that a gamete with $i$ mutations derives from a gamete with $i-1$ mutations is:

$$Q_{i,i-1} = \frac{f_{i-1}UE^{-U}}{f_{i-1}UE^{-U} + f_i e^{-U}} = \frac{f_{i-1}(1-s)^{i-1}UE^{-U}}{e^{-U}Ue^{-U} + f_{i-1}(1-s)^{i-1}e^{-U} + f_{i}(1-s)^{i}e^{-U}}$$

which is a particular case of Equation (3) in (Charlesworth et al. 1995), and $f_i$ is the frequency of the $i$ class, after selection. (I assume here that the mutation rate is sufficiently low that I can neglect mutations from classes other than the adjacent one.)

If the distribution of the frequencies of the classes is close to the deterministic expectation most of the time, then

$$f_0 = e^{-\frac{U}{s}(1-s)}$$

and $f_i = f_{i-1} \frac{U(1-s)/s}{i}$

so that
\[ Q_{i,i-1} = \frac{is}{1+(i-1)s} \]  

(A2)

Suppose that one samples randomly 2 individuals, and that these belong to classes \( i \) and \( j \). If \( i \) and \( j \neq 0 \), there are two possible mutational events in the previous generation: either gamete \( i \) came from the \( i-1 \) class (with probability \( Q_{i,i-1} \)), or gamete \( j \) came from the \( j-1 \) class (with probability \( Q_{j,j-1} \)); if \( i=j \) they can also coalesce, with probability \( 1/Nf_i \), since the size of class \( i \), after selection (which is when we are sampling) is \( Nf_i \). Let \( T_{i,j} \) be the mean time (in generations) back to the common ancestor of a sample of two gametes with \( i \) and \( j \) \((i,j \geq 0)\) mutations. One then has:

\[
T_{i,j} = \left(1 - Q_{i,i-1} - Q_{j,j-1} - \frac{\delta_{ii}}{Nf_i}\right)(T_{i,j} + 1) + Q_{i,i-1}(T_{i-1,j} + 1) + Q_{j,j-1}(T_{i,j-1} + 1) + \frac{\delta_{ii}}{Nf_i}
\]

where \( \delta_{ii} = 1 \) if \( i=j \) and 0 otherwise. Rearranging, one has:

\[
T_{i,j} = \left[ \frac{1}{Q_{i,i-1} + Q_{j,j-1} + \frac{\delta_{ii}}{Nf_i}} \right] \frac{Q_{i,i-1}}{Q_{i,i-1} + \frac{\delta_{ii}}{Nf_i}} T_{i-1,j} + \frac{Q_{j,j-1}}{Q_{i,i-1} + \frac{\delta_{ii}}{Nf_i}} T_{i,j-1}
\]

(A3)

which is equivalent to equation (12) of (Hudson and Kaplan 1994) for a sample size of 2, with the difference that we are counting individuals as post-selection adults. The mean time for the most recent common ancestor of two randomly sampled sequences is then:

\[
T_2 = \sum_{i,j=0}^{m} f_i f_j T_{i,j}
\]

(A4)

and the resulting mean number of pairwise differences relative to the neutral expectation will be:
\[ \frac{k}{k_0}(N,U,s) = \frac{T_k}{N} \]  

(A5)

APPENDIX 4.B

Suppose that one takes a random sample of 2 individuals from a population subject to recurrent mutations with two types of effect, \( s_i \) and \( s_b \), occurring at two different rates, \( U_i \) and \( U_b \), respectively. Suppose one individual carries \( i \) mutations of type \( s_i \) and \( k \) mutations of type \( s_b \) and the other carries \( j \) mutations of type \( s_i \) and \( l \) mutations of type \( s_b \). Let \( T_{i,k,j,l} \) be the time to the most recent common ancestor of these individuals. If the population is close to the deterministic equilibrium this time will be given by:

\[
T_{i,k,j,l} = \left[ \frac{1}{Q_{s_i,i-1} + Q_{b_{i,k-1}} + Q_{s_{j,j-1}} + Q_{b_{l,l-1}} + \frac{\delta_y \delta_{s_i}}{N_f f_k}} \right] + \left[ \frac{Q_{s_i,i-1} + Q_{b_{i,k-1}} + Q_{s_{j,j-1}} + Q_{b_{l,l-1}} + \frac{\delta_y \delta_{s_i}}{N_f f_k}}{Q_{s_i,i-1} + Q_{b_{i,k-1}} + Q_{s_{j,j-1}} + Q_{b_{l,l-1}} + \frac{\delta_y \delta_{s_i}}{N_f f_k}} \right] + \left[ \frac{Q_{s_i,i-1} + Q_{b_{i,k-1}} + Q_{s_{j,j-1}} + Q_{b_{l,l-1}} + \frac{\delta_y \delta_{s_i}}{N_f f_k}}{Q_{s_i,i-1} + Q_{b_{i,k-1}} + Q_{s_{j,j-1}} + Q_{b_{l,l-1}} + \frac{\delta_y \delta_{s_i}}{N_f f_k}} \right]
\]

where

\[
Q_{s_i,i-1} = \frac{is_i}{1 + (i-1)s_i} \quad \text{and} \quad Q_{b_{i,i-1}} = \frac{is_b}{1 + (i-1)s_b}
\]

(B1)

which is the extension of the previous approximation for mutations of equal effects.

When \( n_0^0 < 1 \), because \( Nf_0 < 1 \), I use, as previously, the shifted Poisson distribution with parameter \( \lambda_s = U_s s_i^i - K_s \), where \( K_s = \min\{k: Nf_0^0 f_k \geq 1\} \), so that

\[
Q_{s_i,i-1} = \frac{U_s}{U_s + (1-s_i) \lambda_s}
\]

(B2)
Using these approximations, the mean time to the most recent common ancestor of two random gametes is:

\[ T_2 = \sum_{i,k,l} f_{ik} f_{jl} T_{i,k,l} \text{ and } \frac{k}{k_0} \left( N, U_b, s_b, U_s, s_s \right) = \frac{T_2}{N} \]  

(B3)

with

\[ f_{ij} = \left( \frac{U_z (1-s_z)}{s_z} \right)^i \frac{e^{U_z (1-s_z)}}{i!} \left( \frac{U_b (1-s_b)}{s_b} \right)^j \frac{e^{U_b (1-s_b)}}{j!}. \]
5 Muller's ratchet and the distribution of selection coefficients

5.1 Introduction

Up until now, I have considered two simple ways of modelling Muller's ratchet. This has allowed approximate analytical quantifications of its speed and its effects on variability. I have concluded that, if the vast majority of mutations cause similar, very mildly deleterious, effects on fitness, the process could, in principle, operate even in species with relatively large populations and cause a large reduction in linked neutral variability. I have also concluded that its speed can be considerably accelerated by the presence of another group of strongly deleterious mutations, maintained at equilibrium, and associated with an even larger reduction in variability.

The obvious question one wants to ask now is: how do the results obtained in the previous chapters generalise to a continuous distribution of selection coefficients? Although both the rate of deleterious mutations and the shape of the distribution of their effects on fitness has been the subject of a considerable amount of debate, these are still essentially open questions (Chavarrias et al. 2001; Davies et al. 1999; Deng 1998; Deng et al. 1999; Deng and Lynch 1996; Deng and Lynch 1997; Fry 2001; Keightley 1998; Keightley and Eyre-Walker 1999; Vassilieva et al. 2000). A gamma distribution of mutational effects on fitness is sometimes used both in theoretical studies (Charlesworth 1996a; Keightley 1998) and as a model to fit the data from mutation accumulation experiments (Keightley 1994). Here, I will assume a gamma distribution, since variation of its shape parameter allows the study of very different forms of distributions of selection coefficients. As before, multiplicative fitness is assumed throughout. Although assuming some form of epistasis would probably be closer to what happens in a real biological system, empirical work, at least in some organisms, seems to point towards its form being fluctuating, i.e.
sometimes a new mutation causes more damage when others are present (synergistic epistasis), sometimes less (antagonistic epistasis), such that this effect seems to average out to produce a linear decline in log mean fitness as the number of mutations increases (Elena 1999; Elena and Lenski 1997; Peters and Keightley 2000).

One of the major criticisms of models of the ratchet that assume equal fitness effects of mutations is that they will underestimate its effects. For example, the following has been stated:

"It is sometimes argued that Muller's ratchet does not operate in large populations, but this conclusion is a modelling artefact that occurs when the selection coefficients (s) of all mutations are defined to have the same value (e.g. set equal to the average value of s). When variable selection coefficients are permitted, with a high density of very small selection coefficients (such as those from nonpreferred codons, transposable element inserts and lesions to nonessential genes), then the expected number mutations per genome is >100, causing the expected number of individuals in the highest extant fitness class to be one or a few individuals, and thus Muller's ratchet is expected to be ubiquitous." Rice (1999)

Here I explore how ubiquitous the ratchet is expected to be when considering a continuous distribution of selection coefficients with very different shapes, including one that allows precisely for the small effect mutations to which Rice (1999) refers. I will not consider here mutations for which \( N_s \) is around 1, such as those that are generally considered in models of mutations from preferred to unpreferred codons (see, for example, McVean and Charlesworth (2000) for a treatment of this case), since this case is only realistically modelled by considering reversible mutations, which is inconsistent with a ratchet.

The effect of a continuous distribution of \( s \) is examined by using two kinds of Monte Carlo stochastic simulations. In the first set of simulations, one is only interested on the dependence of speed of Muller's ratchet on the distribution of selection coefficients. The speed is now measured as the rate of decline in log mean fitness per generation. In the second set of simulations, the variability of a neutral
locus embedded in a set of selected loci, subject to the accumulation of deleterious mutations by the ratchet mechanism, is studied. I also compare the values of the Tajima's $D$ statistic (Tajima 1989), obtained with a distribution of $s$, to the results of the constant mutational effect model obtained previously. No analytical approximations are obtained for this case. But the following conclusions will be drawn from the simulation results:

In relation to the speed of the ratchet:

- there is a given combination of the mutation rate, mean value of $s$ and population size, for which the decline in log mean fitness is roughly independent of the distribution of selection coefficients.
- in large populations, assuming constant selection coefficients generally underestimates the speed of the process, as measured by the decline in mean fitness.

In relation to the effects on neutral variability:

- for intermediate values of the mean selection coefficient against the mutations, with values of $U$ for which the ratchet operates, the reduction in linked neutral variability is largely independent of the shape of the distribution of selection coefficients.
- irrespective of the shape parameter of the gamma distribution, negative mean values of Tajima's $D$ are expected if the ratchet operates.

It is in fact the case that, as far as the increase in the mean number of mutations is concerned, the ratchet is expected to be ubiquitous if the shape of the distribution of $s$ is very skewed towards weak mutations (Rice 1999), but this ubiquity is not necessarily reflected on its effects on mean fitness, if the mean selection coefficient and the deleterious mutation rate are small.
5.2 Simulation Methods

5.2.1. Monte Carlo simulations of the speed of the ratchet

Following the simulation work of the chapters 2, 3 and 4, a haploid, non-recombining population of $N$ chromosomes was simulated with the following life cycle: mutation, reproduction and selection. In each generation, mutations to deleterious alleles occur according to a Poisson distribution with mean $U$. The deleterious effects of a mutation ($s$) are drawn from a gamma distribution, with shape parameter $\alpha$ and scale parameter $\beta$, under the constraint that $s > 5/N$. This is to ensure that one is looking at deleterious mutations whose selection coefficient is such that their probability of fixation in a freely recombining environment would be negligible (for example, in the absence of any Hill-Robertson effect, the probability of fixation of a mutation with $Ns = 5$ and $N = 1,000$ is $4.6 \times 10^{-7}$ and with $N = 100,000$ it is $4.5 \times 10^{-9}$ (Kimura 1957)). The vast majority of mutations then stay at low frequencies at any particular locus, so that the assumption of negligible back-mutation may be considered reasonable. The gamma variates were generated according to Dagpunar (1988). Multiplicative fitness effects of the deleterious mutations are assumed. Every generation, the population mean fitness and the numbers of individuals with maximal fitness and with the minimum number of deleterious mutations are recorded.

Populations were run starting from every individual being free of deleterious mutations, for an initial period until the mean fitness achieves its deterministic expectation $\exp(-U)$, which is independent of the distribution of mutational effects under a multiplicative fitness assumption (Johnson 2000; Kimura and Maruyama 1966). After such an initial period, a certain number of generations, more than 10,000 generations and sometimes up to 200,000 generations were run. The rate of decline in population mean fitness was calculated, and the rate of increase in the
minimum number of mutations computed. At least 5 replicates were run for each set of parameters.

**5.2.2 Monte Carlo simulations for neutral variability**

As in chapter 4, for the analysis of the level and pattern of variability, a neutral locus was introduced, in which mutations occur according to the infinite sites model, i.e. each mutation occurs at a new site, at a total rate of $\mu$ at the locus, which has 250 neutral sites. Every $N$ generations, a sample of size $n = 25$ chromosomes is taken from the population, and variability measures at the neutral locus are computed. For each set of parameter values, 10 simulation runs were done, with 5 samples taken per run. A total of 50 samples were used for the computations. The mean number of pairwise differences between randomly sampled sequences, $k$, and the number of segregating sites, $S$, are used as measures of genetic variation in a sample of alleles at the neutral locus. Tajima's $D$ statistic is computed for every sample and the proportion of times that, in the presence of a ratchet, the neutral model is rejected because of very negative $D$s is assessed in the following way. As in chapter 4, for a given $\theta$ value, standard coalescent simulations (Hudson 1990) of the neutral infinite sites mutational model were run and the critical values (at the 95% CI) of the statistic $D$ calculated. Then the results of the forward simulations were used to compute the rejection power, given by the proportion of forward simulations whose observed $D$ was lower than the critical value expected under neutrality: $Pow$.

**5.3. Preliminary considerations**

The effects of an exponential distribution of selection coefficients have been analysed previously by Gessler (1995) and by Butcher (1995), using Monte Carlo
simulations. Gessler (1995) considered a model with constant population size, like the one assumed here. Butcher (1995) considered a model where the population size can change with the accumulation of mutations and the population may go extinct due a mutational meltdown, which I have described in chapter 1. Up to now, no analytical approximation has been proposed for predicting the rate of the ratchet in any of these models. An analytical approximation to the rate of decline in mean population fitness is also not going to be presented here. The problem appears to be one of extreme difficulty.

Application of diffusion theory to calculate the time to extinction of the least-loaded class, as was done in chapter 2 and 3, seems difficult. With a continuous distribution of \( s \), there is no longer a one-to-one correspondence between the number of mutations and the fitness of an individual. I could in principle try to use the equilibrium value of the frequency of the fittest class \( (x_0) \) as the initial condition in the diffusion approximation. This value is given by \( \exp(-U/s_{11}) \) where \( s_{11} \) is the harmonic mean of the distribution of selection coefficients. But to quantify the drift coefficient for solving the diffusion equation, I would have to approximate the decline in mean fitness after the loss of the fittest class. To do this, I have to know the selection coefficient against the mutations in the next fittest class. Because there are now many classes, and due to the lack of correspondence between number of mutations carried by an individual and its fitness, this appears to be difficult. For example, the fitness of an individual with one strongly deleterious mutation may be the same as that with a large number of weakly deleterious mutations. I did not succeed in finding a reasonable approximation to the decline in mean fitness using this approach.

In addition, as pointed out by Gessler (1995), the mutations of smaller effect will preferentially accumulate, and the mutations of stronger effect will be more likely to stay at their equilibrium frequencies. This means that there are going to be
some mutations causing background selection, and others that cause a ratchet. I was unable to find any general way of calculating the fractions causing each of the effects, in order to produce a reasonable approximation for the decline in mean fitness per generation.

Gessler (1995) observed that, in a haploid non-recombining population, the rate of decline in mean fitness was larger with constant selection coefficients than with an exponential distribution, if the mean selection coefficient was smaller than a particular value. Therefore, the assumption of constant $s$ may or may not be conservative regarding the effects of the ratchet. Here, I follow his work to explore different distributions and larger populations, as well as lower mutation rates, which could potentially apply to a proto-Y or neo-Y chromosome, than those studied by Gessler. With the simulation method that I am using to produce the results presented below, I recovered the results obtained in figure 7 of Gessler (1995).

5.4. Results

5.4.1. The ratchet and the decline in mean fitness

In figure 5.1 the various shapes of the distributions of mutational effects considered in the Monte Carlo simulations are shown. For these various shapes I will first ask the question: how does the decline in log mean fitness caused by the ratchet change with increasing population size?

Figures 5.2 and 5.3 show the absolute value of the rate of decline in log mean fitness as a function of the population size, for genomic mutation rates $U = 0.04$ and $U = 0.08$, respectively. The arithmetic mean of $s$ is $E(s) = 0.02$. Note that the log mean fitness is, obviously, declining with time, and that the slope of that decline is plotted on the y-axis of the figures. As expected, for any given shape of the distribution increasing $N$ decreases the speed of the ratchet.
Figure 5.1 The different distributions of selection coefficients with different shape parameters but with the same arithmetic mean $E(s) = 0.02$. $\alpha = 0.5$ for the full line, $\alpha = 1$ (exponential) for the dashed line, $\alpha = 2$ for the full line with triangles and $\alpha = 10$ for the line with crosses.

Figure 5.2 The absolute value of decline in log mean fitness ($\times 10^6$) as a function of population size for different distributions of $s$. $U = 0.04$ and the arithmetic mean value of $s$ is $E(s) = 0.02$. The solid line is the value for a constant mutational model. The other lines are for $\alpha$ values as indicated in the figure. Error bars are not presented in order to make the figure clear, but for $\alpha < 10$ they are within the marker points.
I found that the speed of the ratchet is very similar for an exponential distribution and a distribution with $\alpha < 1$. I also observe that, for a given $U$ and $E(s)$ there is a value of $N, N_i$, say, for which the rate of decline in mean fitness is independent of the shape of the distribution. For populations with sizes above $N_i$, use of a model with constant $s$ underestimates the effects of the ratchet, in terms of
decline in mean fitness. For populations with size below $N$, the reverse is observed. Therefore, in a large population one is likely to greatly underestimate the decline in log mean fitness by assuming a constant mutational model, if indeed the real distribution is close to a gamma distribution with a small value of $\alpha$.

**Figure 5.4.** The absolute value of decline in log mean fitness ($\times 10^6$) as a function of $U$ for different distributions of selection coefficients. $N = 5,000$ and $E(s) = 0.02$. Other symbols as in figure 5.2.

I now ask how the speed of the ratchet changes with an increase in the mutation rate. Figure 5.4 shows an example with $N = 5,000$ and $E(s) = 0.02$. For any given shape of the distribution, decreasing the mutation rate decreases the rate of decline in fitness and this is more pronounced for $\alpha > 1$. Again, I do not observe many differences between an exponential distribution and a distribution with $\alpha < 1$. In figure 5.4, it is shown that, for a given $N$ and average $s$, there is a value of the
genomic mutation rate $U_i$ for which the decline in log mean fitness is largely distribution-independent. Above this value, assuming an equal effects model overestimates the effects of the ratchet; below this value its effects are underestimated.

Figure 5.5 The absolute value of decline in log mean fitness ($x10^6$) as a function of the arithmetic mean of the selection coefficient, for different distributions of selection coefficients. $N = 10,000$ and $U = 0.04$. Other symbols as in the previous figures.

Finally, figure 5.5 shows the same phenomena but now in relation to the mean selection coefficient. For a given $U$ and $N$ there is a value of $s, s_n$, above which the decline in log mean fitness is underestimated by assuming equal selection coefficients, as previously seen in figure 7 of Gessler (1995). Note that if $E(s)$ gets very close to zero, all the curves would tend to zero, since they will tend to the
neutral result. This means that, for very small values of $E(s)$, the effects of the ratchet can become largely irrelevant, irrespective of the distribution of $s$.

The question now becomes: what is the condition that allows us to know if one is underestimating or overestimating the effects of the process by assuming a constant mutational model. I found that, for all the parameter sets tested, the combination of parameters where the decline in log mean fitness does not strongly depend on the distribution of mutational effects (and can therefore be estimated analytically using the approximation proposed in chapter 2, or that proposed by Gessler (1995)) occurs when:

$$n_os_i < 1 \quad (1)$$

where $s_i$ is the arithmetic mean of the selection coefficients. An example of this is provided in table 5.1, where very different parameter combinations are used to produce the condition given above.

One possible interpretation for this condition uses the following reasoning. Assume that the population achieves a state close to equilibrium, so that the size of the fittest class has a value $n_o$. After some time this class will be lost and the ratchet will have clicked. The decline in mean fitness will be given by the difference in mean fitnesses before and after the click. But, with a distribution, the new least-loaded class may be composed of alleles with different fitnesses. If $n_os_i < 1$, then within this class the alleles behave as effectively neutral, such that one of them randomly will increase in frequency to fixation. The change in mean fitness is then approximately $s_i$, i.e. roughly independent of the distribution. On the other hand, if $n_os_i >> 1$, then within the new least-loaded class the alleles are no longer effectively neutral, such that the one with the smaller $s$ will tend to be the one that dominates. Then, the difference in mean fitness will tend to be smaller than expected on a
constant-$s$ model. From the results of chapter 2, this leads to a faster rate of movement of the ratchet.

Table 5.1

Parameter values for which the decline in log mean fitness is approximately independent of the shape of the distribution.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$E(s)$</th>
<th>$n_0$</th>
<th>$n_0E(s)$</th>
<th>$\frac{dLn(w)}{dt}$</th>
<th>2SE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$N=1000$</th>
<th>$U=0.5$</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>4.50</td>
<td>0.111</td>
<td>11</td>
<td>1.23</td>
<td>-2.33x10$^3$</td>
<td>1.1x10$^4$</td>
</tr>
<tr>
<td>1</td>
<td>9.01</td>
<td>0.111</td>
<td>11</td>
<td>1.23</td>
<td>-2.64x10$^3$</td>
<td>2.0x10$^4$</td>
</tr>
<tr>
<td>2</td>
<td>18.02</td>
<td>0.111</td>
<td>11</td>
<td>1.23</td>
<td>-2.71x10$^3$</td>
<td>4.5x10$^5$</td>
</tr>
<tr>
<td>10</td>
<td>90.09</td>
<td>0.111</td>
<td>11</td>
<td>1.23</td>
<td>-2.92x10$^3$</td>
<td>1.1x10$^4$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$N=1000$</th>
<th>$U=0.25$</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>7.81</td>
<td>0.064</td>
<td>20</td>
<td>1.29</td>
<td>-7.98x10$^4$</td>
<td>3.0x10$^5$</td>
</tr>
<tr>
<td>1</td>
<td>15.63</td>
<td>0.064</td>
<td>20</td>
<td>1.29</td>
<td>-8.65x10$^4$</td>
<td>2.2x10$^5$</td>
</tr>
<tr>
<td>2</td>
<td>31.25</td>
<td>0.064</td>
<td>20</td>
<td>1.29</td>
<td>-8.95x10$^4$</td>
<td>7.5x10$^5$</td>
</tr>
<tr>
<td>10</td>
<td>156.25</td>
<td>0.064</td>
<td>20</td>
<td>1.29</td>
<td>-1.01x10$^3$</td>
<td>1.2x10$^4$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$N=1000$</th>
<th>$U=0.05$</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>28.57</td>
<td>0.0175</td>
<td>57</td>
<td>1.01</td>
<td>-5.13x10$^5$</td>
<td>6.4x10$^6$</td>
</tr>
<tr>
<td>1</td>
<td>57.14</td>
<td>0.0175</td>
<td>57</td>
<td>1.01</td>
<td>-5.65x10$^5$</td>
<td>8.6x10$^6$</td>
</tr>
<tr>
<td>2</td>
<td>114.29</td>
<td>0.0175</td>
<td>57</td>
<td>1.01</td>
<td>-6.16x10$^5$</td>
<td>6.8x10$^6$</td>
</tr>
<tr>
<td>10</td>
<td>571.43</td>
<td>0.0175</td>
<td>57</td>
<td>1.01</td>
<td>-6.80x10$^5$</td>
<td>1.1x10$^5$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$N=5000$</th>
<th>$U=0.1$</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>22.73</td>
<td>0.022</td>
<td>53</td>
<td>1.17</td>
<td>-7.40x10$^5$</td>
<td>7.5x10$^6$</td>
</tr>
<tr>
<td>1</td>
<td>45.45</td>
<td>0.022</td>
<td>53</td>
<td>1.17</td>
<td>-7.70x10$^5$</td>
<td>6.4x10$^6$</td>
</tr>
<tr>
<td>2</td>
<td>90.91</td>
<td>0.022</td>
<td>53</td>
<td>1.17</td>
<td>-8.50x10$^5$</td>
<td>3.5x10$^6$</td>
</tr>
<tr>
<td>10</td>
<td>454.55</td>
<td>0.022</td>
<td>53</td>
<td>1.17</td>
<td>-9.50x10$^5$</td>
<td>5.2x10$^6$</td>
</tr>
</tbody>
</table>

$\alpha$ and $\beta$ are the shape and scale parameters of the gamma distribution. $E(s)$ is the arithmetic mean $E(s) = \alpha/\beta$. $n_0$ is calculated for the constant case, $\alpha \to \infty$.

Another way of looking at it is the following: the time between clicks of the ratchet is given by the time to reach equilibrium ($T_{eq}$) plus the time for the stochastic loss of the least-loaded class ($T_{L}$. With a constant $s$ model the major determinant of
is 1/s, and a major determinant of \( T_L \) is \( n_0 \). If \( n_0s >> 1 \), \( T_{eq} << T_L \). Under a variable 
model \( T_{eq} = 1/s_H \) (Johnson 2000) and \( T_L = N \exp(-U/s_H) \) meaning that, as \( s_H \) 
decreases (or as \( \alpha \) decreases), \( T_L \) will decrease faster than \( T_{eq} \) will increase, resulting 
in a much smaller time between clicks than in the constant case.

I should note that the underestimation of the effects of the ratchet in terms of 
decline in mean fitness by an equal effects of mutation model, when \( n_0s >> 1 \), is 
bigger than the overestimation of its effects when \( n_0s << 1 \). For example, in figure 
5.3 when \( N = 10^5, U = 0.08 \) and \( E(s) = 0.02 \), assuming a constant model leads to the 
conclusion that the ratchet does not effectively work, whereas with \( \alpha = 2 \) its 
operation causes the fitness of the population after 500,000 generations to be about 
3% of its initial value.

Since I have followed the decline in population mean fitness and the increase 
in the minimum number of deleterious mutations in the population across 
generations, I can ask what mean selection coefficient, say \( s_w \), is associated with the 
mutations that are causing the observed decline in log mean fitness. The value of \( s_w \) 
will be approximated by:

\[
s_w = \frac{T_{rat} d\ln w}{dt} \tag{2}
\]

where \( T_{rat} \) is the average time in between losses of the least-loaded classes.

In the long term, the rate of increase in the minimum number of deleterious 
mutations (1/\( T_{rat} \)) is the rate of fixation; see the discussion in Charlesworth and 
Charlesworth (1997).

In table 5.2 I show the values of \( T_{rat} \) from the simulations for some sets of 
parameter values. The value of \( s_w \) as calculated by equation (2) is also shown. I 
observed that the mutations that are contributing to the decline in mean fitness are 
those for which \( Ns \) is far below \( NE(s) \) or \( Ns_H \). These mutations have \( Ns_w \) of the order
of 10 to 100, depending on the population size. I then compare the mean time $T_{rot}$ observed in the simulations with the one obtained theoretically by assuming a constant mutational model where the value of $s$ is given by the harmonic mean of the distribution (using the results in chapter 2). As discussed before, I have tried to make this comparison, because the frequency of the fittest class ($x_0$) is given by $\exp(-U/s_H)$ (Johnson 2000).

### Table 5.2

Mean selection coefficient of mutations that cause the decline in fitness

<table>
<thead>
<tr>
<th>$N$</th>
<th>$U$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$E(s)$</th>
<th>$s_H$</th>
<th>$n_0$</th>
<th>$T_{rot}$</th>
<th>$T_{theo}$</th>
<th>$s_w$</th>
<th>$Ns_w$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x10^3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.04</td>
<td>10</td>
<td>500</td>
<td>0.02</td>
<td>0.018</td>
<td>108</td>
<td>564</td>
<td>524</td>
<td>0.014</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>0.04</td>
<td>10</td>
<td>500</td>
<td>0.02</td>
<td>0.018</td>
<td>217</td>
<td>1742</td>
<td>1553</td>
<td>0.014</td>
<td>28</td>
</tr>
<tr>
<td>5</td>
<td>0.04</td>
<td>10</td>
<td>500</td>
<td>0.02</td>
<td>0.018</td>
<td>542</td>
<td>29361</td>
<td>29496</td>
<td>0.0106</td>
<td>53</td>
</tr>
<tr>
<td>1</td>
<td>0.04</td>
<td>2</td>
<td>100</td>
<td>0.02</td>
<td>0.015</td>
<td>69</td>
<td>355</td>
<td>318</td>
<td>0.009</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>0.04</td>
<td>2</td>
<td>100</td>
<td>0.02</td>
<td>0.011</td>
<td>131</td>
<td>708</td>
<td>626</td>
<td>0.0037</td>
<td>18</td>
</tr>
<tr>
<td>10</td>
<td>0.04</td>
<td>2</td>
<td>100</td>
<td>0.02</td>
<td>0.01</td>
<td>223</td>
<td>1167</td>
<td>998</td>
<td>0.0027</td>
<td>27</td>
</tr>
<tr>
<td>100</td>
<td>0.04</td>
<td>2</td>
<td>100</td>
<td>0.02</td>
<td>0.01</td>
<td>1919</td>
<td>14034</td>
<td>1.7x10^7</td>
<td>0.00065</td>
<td>65</td>
</tr>
<tr>
<td>5</td>
<td>0.04</td>
<td>1</td>
<td>50</td>
<td>0.02</td>
<td>0.0077</td>
<td>28</td>
<td>310</td>
<td>215</td>
<td>0.0035</td>
<td>17</td>
</tr>
<tr>
<td>10</td>
<td>0.04</td>
<td>1</td>
<td>50</td>
<td>0.02</td>
<td>0.006</td>
<td>16</td>
<td>410</td>
<td>192</td>
<td>0.0028</td>
<td>28</td>
</tr>
<tr>
<td>100</td>
<td>0.04</td>
<td>1</td>
<td>50</td>
<td>0.02</td>
<td>0.004</td>
<td>2</td>
<td>665</td>
<td>237</td>
<td>0.00098</td>
<td>98</td>
</tr>
<tr>
<td>10</td>
<td>0.04</td>
<td>0.5</td>
<td>25</td>
<td>0.02</td>
<td>0.0043</td>
<td>1</td>
<td>226</td>
<td>200</td>
<td>0.0015</td>
<td>15</td>
</tr>
<tr>
<td>1</td>
<td>0.08</td>
<td>10</td>
<td>500</td>
<td>0.02</td>
<td>0.018</td>
<td>12</td>
<td>99</td>
<td>89</td>
<td>0.015</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>0.08</td>
<td>10</td>
<td>500</td>
<td>0.02</td>
<td>0.018</td>
<td>23</td>
<td>151</td>
<td>135</td>
<td>0.015</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>0.08</td>
<td>10</td>
<td>500</td>
<td>0.02</td>
<td>0.018</td>
<td>59</td>
<td>308</td>
<td>282</td>
<td>0.0144</td>
<td>72</td>
</tr>
</tbody>
</table>

$s_H$ is the harmonic mean of the selection coefficient. $T_{rot}$ is the mean time to loss of the class with the minimum number of mutations. $T_{theo}$ is the mean time to loss of the least loaded class calculated using the diffusion approximation in chapter 2, but with $s = s_H$. $s_w$ is calculated using the decline in log mean fitness and $T_{rot}$ as explained in the text. Other parameters are as in table 5.1.

Only when $\alpha$ is large does the theoretical calculation agree reasonably well with the simulation result. From the results in Table 5.2, I find that, if the real distribution is close to an exponential distribution, the results using a constant model
with $s = s_\mu$, will tend to overpredict the number of fixations expected to occur over a large period of evolutionary time, in large populations. But with $\alpha = 2$ and $N = 100,000$ I find that the opposite occurs.

5.4.2. The ratchet and the level of diversity

I now look at the effects of the ratchet on the levels and patterns of neutral variability, under a distribution of $s$. Due to the amount of computer time required for these simulations, and because I cannot run coalescent simulations for these mutational models, only relatively small populations will be considered.

In chapter 4, I found that, in conditions favourable to the operation of the ratchet, for a given value of population size and mutation rate, there was a wide range of values of the selection coefficient (intermediate values) for which the reduction in variability levels was approximately unchanged. Given the observation that, under the movement of the ratchet, the level of variation is much more sensitive to changes in $U$ and $N$ than in $s$, one expects that the shape of the distribution will also not change the level of variability set by a particular value of $N$ and $U$, as long as the ratchet operates. In Figures 5.6 and 5.9, I try to show some examples of this expectation.

In figure 5.6 the population size is set to a constant value, of 6,000 chromosomes, and the arithmetic mean of $s$ has the value 0.02. As the value of $\alpha$ changes, the shape of the distribution changes from being very skewed towards a high density of small effect mutations to a rather more symmetrical one with a larger number of mutations of intermediate effect. Three values of the mutation rate are considered.

For $U = 0.1$, there is both a considerable decline in log mean fitness and a high rate of loss of the class with the least number of mutations, which in the long
run will parallel the rate of fixation. Both for $U = 0.1$ and $U = 0.05$, the reduction in the mean number of pairwise differences, $k / k_0$, is largely independent of the shape of the distribution. This suggests that we can use the constant mutational model to set the approximate level of variability that is expected for a given $U$ and $N$. The theoretical prediction applying the coalescent approximation from chapter 4 with a constant selection coefficient of 0.02 (which is the arithmetic mean of $s$ for every shape of the distribution) is 0.031 for $U = 0.1$ and 0.086 for $U = 0.05$.

![Figure 5.6](image.png)

**Figure 5.6** The reduction in neutral diversity, $k/k_0$, for different values of $U$ and different distributions of selection coefficients with the same arithmetic mean. $N = 6000$, and $E(s) = 0.02$. $\theta = 6$.

For $U = 0.1$, this value is slightly below the values observed in the simulations with variable selection coefficients. The rate of decline in log mean fitness with $\alpha = 5$ is $9.5 \times 10^{-5}$ and the mean time between losses of the class with the
minimum number of mutations \((T_{\text{mu}})\) is around 115 generations; with \(\alpha = 0.5\) the values are \(6.4 \times 10^5\) and 50 respectively.

For a much smaller mutation rate of 0.025, there is effectively no ratchet with \(\alpha = 5\), but with \(\alpha = 0.5\), I observe a decline in log mean fitness of \(3 \times 10^6\) with \(T_{\text{rot}}\) of around 600 generations. In accordance with these observations, for \(\alpha = 5\), the population is essentially at the mutation–selection deterministic equilibrium and the reduction in genetic diversity is very close to that expected by calculating \(f_0\), using the harmonic mean of \(s\). Such a calculation results in a reduction in the mean number of pairwise differences of 0.215. For \(\alpha = 0.5\), this equilibrium is being lost and the level of variability is higher than the predicted by such a calculation (which is 0.012), as expected from the results in chapter 4.

For \(U = 0.05\), I observe clicks of the ratchet for every value of \(\alpha\) considered in figure 5.6. The predicted reduction in variability as \(\alpha\) tends to infinity (constant \(s\) model) is 0.086. This value is slightly higher than the observed for \(\alpha = 5\), \(k/k_0 \pm 2\text{SE} = 0.072 \pm 0.017\), and slightly lower than the observed for \(\alpha = 0.5\), \(k/k_0 \pm 2\text{SE} = 0.118 \pm 0.023\).

In chapter 4, I showed that, under a constant \(s\) model, the approximation that a non-recombining population is equivalent to a structured population constituted by demes of different sizes is reasonable for predicting levels of variability. The sizes of the demes were set by the values of \(N\), \(U\) and \(s\). I concluded that, as soon as the ratchet starts to operate, the time to coalesce within the least loaded class \((n_0)\) becomes smaller than the time to migrate into that class \((-1/s)\), and therefore the reduction in diversity is overestimated by \(f_0\). Similarly, with a gamma distribution I also observe the same overestimation, as soon as the ratchet starts clicking.

Trying to obtain a quantification of the reduction in diversity using the coalescent approach is now more complicated. A class of individuals with a given fitness does not simply correspond to the number of mutations it contains, and in
addition there are many different fitness classes. I can try to approximate the class sizes by grouping the individuals into discrete classes, \( i \). I can assume the fitness difference between class \( i \) and \( i-1 \) to be approximately \((1-s_{H})\), where \( s_{H} \) is the harmonic mean of the distribution. The harmonic mean is chosen because, for an arbitrary distribution of selection coefficients, the frequency of the fittest class at equilibrium is \( \exp(-U/s_{H}) \) (Johnson 2000).

![Graph](image.png)

**Figure 5.7.** The discrete distribution of fitness classes from the simulations of a gamma distribution with \( \alpha = 5 \) (labelled simulated) and that expected at deterministic equilibrium under a constant model with \( s = s_{H} \) (labelled Poisson).

In figure 5.7 I plot the sizes of these discrete classes obtained from the simulations, for \( \alpha = 5 \). I compare them with a Poisson distribution with mean \( \theta_{H} = U/s_{H} \) (which is the constant s model result with \( s = s_{H} \)). The value of \( N \) is 6,000; \( U = 0.05 \), the arithmetic mean value of \( s \) is 0.02, and the harmonic mean value of \( s \) is 0.016. As can be seen, the differences between the mean class size observed in the
simulations and the Poisson distribution are not very big for these parameter values. The observed mean number of pairwise differences, relative to the neutral expectation, is 0.072, and that calculated using the coalescent approach proposed in chapter 4, assuming that the Poisson distribution for class size set by $s_H$ is 0.07 (note that the value of the frequency of the least-loaded class in an infinitely large population $f_0$ is in this case 0.046). The agreement between the simulations and the approximate theoretical value seems reasonably good with this value of $\alpha$.

Figure 5.8. The discrete distribution of fitness classes from the simulations of a gamma distribution with $\alpha = 0.5$ (labelled simulated) and a shifted Poisson distribution as assumed in a constant model with $s = s_H$ (labelled Shifted Poisson).

But for $\alpha = 0.5$, applying the results of chapter 4 with $s = s_H$ underestimates the value of $k / k_0$ observed in the simulations. The value observed in the simulations is 0.118 but that obtained by using the harmonic mean in the coalescent approach results in a value of 0.066 (note that now $f_0$ is 0.00013). In figure 5.8, I compare the discrete distribution obtained from the simulations and that assumed in the coalescent
approximation. Clearly in this case the distributions are very different, which may explain the lack of agreement between the observed and predicted reduction in diversity using this approach.

I therefore conclude that this approach only produces reasonable results when the ratchet is operating very slowly.

![Graph](image)

**Figure 5.9** The mean reduction in neutral diversity, $k/k_0$, as a function of $N$ and different distributions of selection coefficients with the same arithmetic mean. $U=0.05$ and $E(s) = 0.02$.

In figure 5.9 the effect of $N$ is analysed. Clearly, for large $\alpha$, increasing $N$ decreases the level of variability. But this decrease reaches saturation after a certain value of $N$. This is similar to what I have observed when I studied the constant $s$ model in chapter 4 (see figure 4.2). After a certain value of $N$, the ratchet stops operating (or operates at a very low speed), and the background selection result is
expected. For example, for \( N = 9000 \), I did not observe the effects of the ratchet in the simulations with large \( \alpha \), and the observed reduction in the level of diversity is very close to that expected by calculating \( f_0 \) (which is independent of \( N \)). For small \( \alpha \), increasing \( N \) also decreases the level of variability but the value of \( N \) for which such saturation occurs is higher than in the previous case. One can interpret a decrease in \( \alpha \) as a decrease in the strength of selection, since the smaller the \( \alpha \), the larger the fraction of weaker mutations. This effect is similar to that observed in figure 4.1 (see chapter 4): for small values of \( s \) (which correspond to small \( \alpha \) in figure 5.9), increasing \( N \) increases the reduction in the mean number of pairwise differences, \( k/k_0 \).

Overall, the simulations suggest that for intermediate values of the mean selection coefficient and a somewhat bell-shaped distribution of \( s \), the reduction in the level of diversity is essentially set by \( N \) and \( U \), i.e., it is largely independent of \( \alpha \) and \( E(s) \), as long as the ratchet operates. Under this circumstances it will be close to that assumed by a constant \( s \) model. For an exponential distribution or with \( \alpha < 1 \), the levels of diversity can be slightly higher than those calculated by assuming an equal effects model. Figure 5.6 and 5.9 suggest that, for sufficiently large values of \( N \) and \( U \), the level of variability is roughly independent of the shape of the distribution.

In tables 5.3 and 5.4 I show the average values of Tajima’s \( D \) corresponding to the parameter sets in figures 5.6 and 5.9. I also show the power to reject neutrality calculated as explained in the methods. As observed previously in chapter 4, if the ratchet operates, negative values of Tajima’s \( D \) are expected irrespective of the shape of the distribution. For the \( \theta \) value considered, average Tajima’s \( D \) are close to \(-1\). There is also some power to reject neutrality, as found in chapter 4. The results in table 5.4 also suggest that as \( N \) becomes large, there is more power for small values of \( \alpha \).
Table 5.3

Average Tajima’s $D$ and Power for $N = 6,000$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$D$</th>
<th>2SE</th>
<th>Pow</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>U=0.025</td>
</tr>
<tr>
<td>0.5</td>
<td>25</td>
<td>-0.99</td>
<td>0.18</td>
<td>20</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>-0.54</td>
<td>0.22</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>-0.69</td>
<td>0.24</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>250</td>
<td>-0.41</td>
<td>0.25</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>U=0.05</td>
</tr>
<tr>
<td>0.5</td>
<td>25</td>
<td>-0.88</td>
<td>0.21</td>
<td>29</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>-0.87</td>
<td>0.17</td>
<td>17</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>-0.83</td>
<td>0.23</td>
<td>18</td>
</tr>
<tr>
<td>5</td>
<td>250</td>
<td>-0.94</td>
<td>0.19</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>U=0.1</td>
</tr>
<tr>
<td>0.5</td>
<td>25</td>
<td>-0.99</td>
<td>0.17</td>
<td>27</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>-0.82</td>
<td>0.22</td>
<td>25</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>-0.88</td>
<td>0.19</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>250</td>
<td>-0.69</td>
<td>0.25</td>
<td>19</td>
</tr>
</tbody>
</table>

$\alpha$ and $\beta$ are the parameters of the gamma distribution. SE means standard error. $\theta = 6$ and sample size is 25. $Pow$ is the power to reject neutrality assuming knowledge of $\theta$.

Table 5.4

Average Tajima’s $D$ and Power for $U=0.05$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$D$</th>
<th>2SE</th>
<th>Pow</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$N = 3000$</td>
</tr>
<tr>
<td>0.5</td>
<td>25</td>
<td>-0.83</td>
<td>0.20</td>
<td>22</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>-0.85</td>
<td>0.22</td>
<td>22</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>-0.88</td>
<td>0.23</td>
<td>28</td>
</tr>
<tr>
<td>5</td>
<td>250</td>
<td>-0.93</td>
<td>0.21</td>
<td>30</td>
</tr>
<tr>
<td>100</td>
<td>5000</td>
<td>-0.96</td>
<td>0.21</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$N = 9000$</td>
</tr>
<tr>
<td>0.5</td>
<td>25</td>
<td>-1.14</td>
<td>0.17</td>
<td>41</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>-0.94</td>
<td>0.20</td>
<td>25</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>-0.74</td>
<td>0.25</td>
<td>26</td>
</tr>
<tr>
<td>5</td>
<td>250</td>
<td>-0.75</td>
<td>0.25</td>
<td>23</td>
</tr>
<tr>
<td>100</td>
<td>5000</td>
<td>-0.51</td>
<td>0.24</td>
<td>4</td>
</tr>
</tbody>
</table>

Parameters are as in table 5.3.
5.5. Discussion

Muller’s ratchet is classically studied under the assumption of equal fitness effects of all deleterious mutations. Under this assumption, there is a one-to-one correspondence between the fitness of an individual and the number of mutations that it carries. This implies that the effects of its operation on the decline in fitness and on neutral variability can be quantified by assuming that the population is divided into various well-defined fitness classes. I have considered this model in chapter 2 and chapter 4 solely for its mathematical simplicity. But to move towards a better understanding of the potential biological effects of the ratchet, one must drop such an unrealistic assumption. In chapter 3, I studied a two-class of mutations model to account for the simultaneous presence of mutations with very different selection coefficients. Here I have studied different models of the distribution of selection coefficients. I have only considered a gamma distribution and it is possible that the real distribution would be bimodal (Davies et al. 1999; Keightley and Eyre-Walker 1999). But since a gamma distribution allows the study of many different shapes, it can be considered as an initial step towards comparing a variable selection coefficient model with a constant one. The simulation results presented here suggest certain conclusions that I will discuss below. I should nevertheless point out that more simulation and theoretical work is needed in order to get to a better understanding and quantification of this process under more general conditions, although I feel that this will be a hard task.

5.5.1 Speed of Muller’s ratchet

Under an equal effect of mutations model, I have previously quantified the speed of the ratchet as the inverse of the mean time between its clicks (chapter 2 and 3). The speed of the ratchet is also the rate of fixation of deleterious alleles
I then calculated the decline in log mean fitness by simply multiplying this value by $s$. Under a continuous distribution the comparison is more subtle, since it requires knowledge of the selection coefficient of the mutations that are actually accumulating. I now define the speed of the ratchet as the rate of decline in log mean fitness itself.

I have found that there is a critical condition, $n_0s \sim 1$, for which the speed of the ratchet is largely independent of the shape of the distribution. If one assumes a constant $s$ model then, when $n_0s >> 1$, the speed of the ratchet is underestimated, and when $n_0s << 1$ it is overestimated. When $n_0s \sim 1$, the speed of the ratchet can be approximated analytically by using the results of chapter 2. I also found that for large populations, with a moderate deleterious mutation rate the effects of the ratchet can be greatly underestimated by assuming constant $s$. For the same $N$, $U$ and mean $s$ I did not find significant differences in the speed of the ratchet between an exponential distribution and a distribution with shape parameter $< 1$. As is intuitively obvious for the same $N$, $U$ and mean $s$, the smaller the shape parameter of the gamma distribution the smaller the time between losses of the least mutated class.

### 5.5.2 Muller's ratchet and neutral variation

I have found that, for intermediate values of the mean selection coefficient, the reduction in the level of neutral variability, as measured by the mean number of pairwise differences in a sample of alleles, is approximately independent of the shape of the distribution, especially if the mutation rate and the population size are not too small.

While in an equilibrium non-recombining population under recurrent mutation to deleterious alleles, the level of variability is determined by the value of the harmonic mean of $s$, $s_H$ (Charlesworth 1996a), in a non-equilibrium population
driven by the ratchet, it is not. If it were determined by $s_H$ then, as the shape parameter of the distribution gets smaller, the level of variation would get smaller, which I do not observe (see figure 5.9). When I compared the expected reduction in variability calculated using the coalescent approach under a constant $s$ model (whose value of $s$ is $s_H$) with that obtained in the simulations, I found that this approach would only produce reasonable results when $\alpha$ was large, which is when the harmonic mean is close to the arithmetic mean.

### 5.5.3 Muller's ratchet and the Y chromosome

Given the conclusions summarised above, I now want to try and extrapolate the results in relation to possible parameters for an evolving Y or neo-Y chromosome. I will tentatively discuss how one expects the conclusions drawn in chapter 4 in relation to the degeneration of Y chromosomes to be changed under a continuous distribution. Since I have not produced any analytical results in this study I can only state qualitative tentative expectations. I do this in table 5.5. As I mentioned in chapter 4, $N = 125,000$ is a possible reasonable value for the neo-Y chromosome of *D. miranda* (Yi 2000), and $N = 500,000$ for the *Silene latifolia* Y chromosome (Filatov et al. 2000).

I should point out that, although I have run simulations for the speed of the ratchet with large population sizes, I did not do so for the effects on the level and pattern of variability. Such simulations would be needed to validate the expectations on table 5.5. Perhaps a quick way of doing this is the following. If one is simply interested on the extent to which the level of neutral diversity will be reduced in a large population, one can try to run forward simulations in which the neutral locus is modelled according to the infinite alleles model. This will not allow the calculation
of Tajima’s $D$, but will shed some light on the reduction in diversity, and will reduce computing time.

I have found that, although when $n_o s < 1$ the rate of decline in fitness is overestimated by a constant $s$ model, such an overestimation is small. One can then take the values of $w/w_i$ in table 5.5 (those below the dashed lines) as good approximations to those under a continuous distribution. But since, when $n_o s > 1$, the underestimation of $w/w_i$ can be substantial (values above the dashed line in table 5.5), I have run simulations for $N = 125,000$, $U = 0.01$ and $E(s) = 0.01$ and $E(s) = 0.004$, with an exponential distribution. For $E(s) = 0.004$, after 100,000 generations (after the mean fitness had achieved the expected equilibrium value of $\exp(-U) = 0.99$) I observed 75 clicks of the ratchet, and the population mean fitness was 0.967 after this time. For $E(s) = 0.01$, I observed 14 clicks and the population mean fitness was 0.988. For this parameter set the ratchet seems unable to cause a considerable decline in mean fitness even under an exponential distribution. With $N = 500,000$, $U = 0.01$ and $E(s) = 0.01$, no ratchet is expected in any reasonable amount of evolutionary time.

In conclusion from table 5.5, the ratchet is expected to be important in shaping the evolution of $Y$ or neo-$Y$ chromosomes if the average selection coefficient is below 1% (or the mutation rate is well above what I assumed). A considerable reduction in neutral diversity is expected due to its operation.
Table 5.5

Expectations for a hypothetical Y chromosome under a continuous distribution

<table>
<thead>
<tr>
<th></th>
<th>constant s model</th>
<th>Continuous distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Drosophila miranda neo-Y</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U = 0.01$</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>$N = 125,000$</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$s$</td>
<td>$k/k_0$</td>
<td>No. of fixations</td>
</tr>
<tr>
<td>0.01</td>
<td>0.4</td>
<td>0</td>
</tr>
<tr>
<td>0.005</td>
<td>0.14</td>
<td>0</td>
</tr>
<tr>
<td>0.004</td>
<td>0.087</td>
<td>0</td>
</tr>
<tr>
<td>0.002</td>
<td>0.018</td>
<td>121</td>
</tr>
<tr>
<td>0.001</td>
<td>0.021</td>
<td>570</td>
</tr>
<tr>
<td>0.0005</td>
<td>0.0283</td>
<td>750</td>
</tr>
<tr>
<td>0.0002</td>
<td>0.042</td>
<td>1502</td>
</tr>
<tr>
<td><strong>$U = 0.03$</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>$N = 125,000$</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>0.05</td>
<td>0</td>
</tr>
<tr>
<td>0.005</td>
<td>0.0073</td>
<td>330</td>
</tr>
<tr>
<td>0.004</td>
<td>0.0068</td>
<td>998</td>
</tr>
<tr>
<td>0.002</td>
<td>0.009</td>
<td>1000</td>
</tr>
<tr>
<td>0.001</td>
<td>0.012</td>
<td>2994</td>
</tr>
<tr>
<td>0.0005</td>
<td>0.016</td>
<td>5495</td>
</tr>
<tr>
<td>0.0002</td>
<td>0.021</td>
<td>8197</td>
</tr>
<tr>
<td><strong>Silene latfolia Y</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U = 0.01$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N = 500,000$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>0.37</td>
<td>0</td>
</tr>
<tr>
<td>0.005</td>
<td>0.137</td>
<td>0</td>
</tr>
<tr>
<td>0.004</td>
<td>0.084</td>
<td>0</td>
</tr>
<tr>
<td>0.002</td>
<td>0.0096</td>
<td>9</td>
</tr>
<tr>
<td>0.001</td>
<td>0.0059</td>
<td>514</td>
</tr>
<tr>
<td>0.0005</td>
<td>0.008</td>
<td>1000</td>
</tr>
<tr>
<td>0.0002</td>
<td>0.012</td>
<td>1400</td>
</tr>
<tr>
<td><strong>$U = 0.03$</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>$N = 500,000$</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>0.05</td>
<td>0</td>
</tr>
<tr>
<td>0.005</td>
<td>0.004</td>
<td>29</td>
</tr>
<tr>
<td>0.004</td>
<td>0.0022</td>
<td>367</td>
</tr>
<tr>
<td>0.002</td>
<td>0.003</td>
<td>1000</td>
</tr>
<tr>
<td>0.001</td>
<td>0.0033</td>
<td>2500</td>
</tr>
<tr>
<td>0.0002</td>
<td>0.0055</td>
<td>7692</td>
</tr>
</tbody>
</table>

$w/w_i$ is the ratio of the expected mean fitness after 500,000 generations to initial mean fitness. $k/k_0$ is the reduction in the mean number of pairwise differences. + means that the quantity will tend to be higher than in the constant $s$ model; - means that the quantity will tend to be lower than in the constant $s$ model; ~ means that the quantity is expected to be similar to that in the constant $s$ model. Below the dashed box differences in $w/w_i$ are not expected to be large.
Bibliography


Appendix of published papers
The Degeneration of Asexual Haploid Populations and the Speed of Muller’s Ratchet

Isabel Gordo and Brian Charlesworth

Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom

Manuscript received August 23, 1999
Accepted for publication November 22, 1999

ABSTRACT

The accumulation of deleterious mutations due to the process known as Muller’s ratchet can lead to the degeneration of nonrecombining populations. We present an analytical approximation for the rate at which this process is expected to occur in a haploid population. The approximation is based on a diffusion equation and is valid when \(N \exp(-u/s) \gg 1\), where \(N\) is the population size, \(u\) is the rate at which deleterious mutations occur, and \(s\) is the effect of each mutation on fitness. Simulation results are presented to show that the approximation estimates the rate of the process better than previous approximations for values of mutation rates and selection coefficients that are compatible with the biological data. Under certain conditions, the ratchet can turn at a biologically significant rate when the deterministic equilibrium number of individuals free of mutations is substantially \(>100\). The relevance of this process for the degeneration of Y or neo-Y chromosomes is discussed.

New mutations arise continuously within populations, the vast majority probably being slightly deleterious (Crow 1993). An effectively infinite asexual population subject to recurrent deleterious mutations will achieve an equilibrium resulting from the continuous appearance of new mutations opposed by selection against them: deterministic mutation-selection balance. With independent and identical effects of each mutation, the equilibrium number of mutations in a haploid randomly mating population follows a Poisson distribution with mean \(u/s\) (Kimura and Maruyama 1966; Haigh 1978), where \(u\) is the per genome mutation rate and \(s\) is the selection coefficient against a deleterious mutation. At this deterministic equilibrium, the number of individuals free of mutations (the best or "least-loaded" class) in a large population of \(N\) breeding adults, is \(n_0 = N \exp(-u/s)\).

But random genetic drift plays a role in a finite population, and it may perturb this equilibrium, leading to the loss of the best class. In the absence of recombination, and with the reasonable assumption that back mutation is negligible for strongly selected mutations, the loss of this class is irreversible and mutations will continually accumulate in the population, leading to the decline of its mean fitness. This is the process known as Muller’s ratchet (Muller 1964; Felsenstein 1974). Once the best class is lost, the new least-loaded class is now the one that has one mutation, but this is also subject to stochastic loss, so that a repetition of successive losses of the least-loaded class can be seen as successive clicks of the ratchet.

One of the important questions in this process concerns the rate or speed at which it operates, or: how much time does it take for the ratchet to click one notch? Because the degeneration of the Y chromosome (Charlesworth 1978, 1996; Rice 1994; Charlesworth and Charlesworth 1997, 1998) and the fate of asexual populations (Pamilo et al. 1987; Gabriel et al. 1993; Lynch et al. 1993, 1995) may involve the operation of Muller’s ratchet, the quantification of its rate is of great biological importance.

Haigh (1978) suggested that the most important parameter for the ratchet mechanism would be \(n_0\), because it is the loss of the best class that drives the process; the smaller the \(n_0\), the faster the ratchet is likely to be. Although Haigh suggested an expression [Equation 9a in Haigh (1978)] for the time between clicks of the ratchet, this is basically a fit to his simulation results. Bell (1988) also suggested an expression based on a fit of the time to extinction of the best class as a function of \(n_0\) (roughly \(10n_0\)). Several attempts toward the quantification of the process have subsequently been made, either using a quantitative genetics approach for estimating the rate of change of the average number of mutations (Pamilo et al. 1987; Gabriel et al. 1993; Lynch et al. 1993; Higgs and Woodcock 1995; Prügel-Bennett 1997) or using diffusion theory to calculate the mean time to loss of the least-loaded class (Stephan et al. 1993; Charlesworth and Charlesworth 1997).

In this article, it is shown that the size of the best class is not sufficient to predict the speed of the ratchet. For the same value of \(n_0\), the ratchet can turn at very different speeds. An approximation based on a diffusion...
equation, for the case $n_0 > 1$, which almost certainly applies to the evolution of the Y chromosome (Charlesworth 1996), is presented, together with simulations of sexual haploid populations to check its validity. Gessler (1995) has derived an approximation that seems to work well for the case $n_0 < 1$. Comparisons between the simulation results and the predictions from the analytical expression suggest that the formulas seem to predict the rate of the process better than the previous formulas for a region of parameter space that is of biological interest (Charlesworth and Charlesworth 1997).

**APPROXIMATION BASED ON THE Diffusion Equation**

We start with a haploid asexual population at equilibrium under mutation-selection balance, with $x_0 = \exp(-u/s)$ being the frequency of individuals in the least-loaded class. The existence of this equilibrium requires that the number of individuals in the best class ($n_0 = Nn_0$) be $>1$. When $n_0 < 1$, this equilibrium may not be approached in a finite population, because it requires the existence of individuals that have a very low probability of actually being present (Gessler 1995; Gessler and Xu 1999).

The way the frequency of the best class varies through time is dictated by mutation taking it below the equilibrium value, selection restoring it to that value, and by stochastic fluctuations due to drift. We wish to quantify how much time it takes for the frequency of the best class (from now on denoted by $x$), with initial value $x_0$, to reach the value zero. One way to do this is to use a diffusion equation for the density function of the time until absorption occurs, subject to the condition that $x = 0$ is the only absorbing state (Ewens 1979, Equations 4.39, 4.40, p. 123). To solve this equation, one has to evaluate the deterministic change (drift coefficient) and stochastic variance (diffusion coefficient) in $x$.

Assuming a Wright-Fisher population, the diffusion coefficient is just the variance due to binomial sampling of $N$ individuals from the previous generation (Stephan et al. 1993; Charlesworth and Charlesworth 1997),

$$b(x) = \frac{x(1-x)}{N} = \frac{x}{N} \tag{1a}$$

assuming $x \ll 1$.

The drift coefficient, representing the expected change in $x$ due to mutation and selection (Stephan et al. 1993; Charlesworth and Charlesworth 1997), is

$$a(x) = \frac{x(e^{-s} - \bar{w})}{\bar{w}} = \Delta \bar{w} \frac{x}{\bar{w}} \tag{1b}$$

where $\bar{w}$ is the mean fitness of the population and $\Delta \bar{w}$ is the difference between the current mean fitness and the mean fitness at equilibrium.

Let us now make the simplifying assumption that the changes in mean fitness are sufficiently small that they can be approximated by small perturbations from the equilibrium value (Stephan et al. 1993). This implies that the system is close to its equilibrium state most of the time, as is supported by our simulations (see Figure 9). We assume that the perturbations in $\bar{w}$ are mostly due to small fluctuations in the least-loaded class. This assumption was employed by Stephan et al. (1993) and Charlesworth and Charlesworth (1997), and is justified in practice by the observation that, for large $N$, the distribution among the classes that are present at any time remains close to the Poisson distribution given by the deterministic equilibrium formula (see Table 1 and Figure 4 of Charlesworth and Charlesworth (1997)). We may express the mean fitness close to equilibrium as a Taylor expansion in $x/x_0$:

$$\Delta \bar{w} = K \left(1 - \frac{x}{x_0}\right), \tag{2b}$$

as previously assumed by Stephan et al. (1993), where $K = x_0 [\delta \bar{w}/\delta x]_0$.

When the frequency $x$ is above its equilibrium value, the system responds with a reduction of mean fitness ($\Delta \bar{w} < 0$) toward the equilibrium value, and when the opposite happens, so that $x$ goes below $x_0$ ($\Delta \bar{w} > 0$), the system responds with an increase in $\bar{w}$. The forces underlying the response of the system toward equilibrium are selection (when $x < x_0$) and mutation (when $x > x_0$), which are parameterized by $K$ in our small-perturbation model. $K$ can be estimated as follows. If by chance the least-loaded class goes extinct ($x = 0$), then the ratchet has clicked and the mean fitness will decline toward a new deterministic equilibrium value, $(1 - s) e^{-s}$. Then the net loss of mean fitness due to a click of the ratchet would be given by $\Delta \bar{w} = se^{-s}$ if the distribution instantaneously recovered its Poisson equilibrium with a new least-loaded class with abundance $n_0$, after the loss of the least-loaded class. In practice, stochastic fluctuations mean that this equilibrium is never achieved exactly.

In fact, the approach to the neighborhood of equilibrium takes some time (Haigh 1978), and just after a click the new best class is above its equilibrium value, so that $\Delta \bar{w} < se^{-s}$. Haigh (1978) showed that, after a click, the new least-loaded class rapidly approaches a value close to $1.6 n_0$, and then the approach to the new equilibrium value is slower. If this is the case, Equation 2b means that the reduction in mean fitness after a click
will not be $se^{-x}$ but $\sim 0.6se^{-x}$. We thus set $K \sim 0.6se^{-x}$ and we test how accurate this approximation is with the help of simulations (see simulation results below).

We now may write the drift coefficient as

$$a(x) = 0.6s(1 - x/x_0)x.$$  \hspace{1cm} (2c)

Using these drift and diffusion coefficients, the time spent in the frequency interval $[0, x_0]$ (Ewens 1979, Equation 4.39) is

$$T_{0,x_0} = \int_0^{x_0} \frac{2N}{x} G(x) \left[ \int_0^x G(x') \, dx' \right] \, dx$$  \hspace{1cm} (3a)

and the time spent in the interval $[x_0, 1]$ is (Ewens 1979, Equation 4.40)

$$T_{x_0,1} = \int_{x_0}^1 \frac{2N}{x} G(x) \left[ \int_0^{x_0} G(x') \, dx' \right] \, dx.$$  \hspace{1cm} (3b)

where

$$G(x) = \exp \left[ -2 \frac{a(x)}{b(x)} \right] = \exp \left[ \frac{2N0.6s}{x_0} \left( \frac{1}{x} - x_0 \right) \right]$$

Using expressions 3a and 3b, and evaluating the integrals numerically for a given population size, mutation rate, and selection coefficient, we obtain the expected time to loss of the least-loaded class as

$$T(N, u, s) = T_{0,x_0} + T_{x_0,1}.$$  \hspace{1cm} \hspace{1cm}

SIMULATION METHODS

For a given population size ($N$), genomic mutation rate to deleterious mutations ($u$), and selection coefficient against each mutation ($s$), haploid asexual populations were simulated starting at mutation-selection equilibrium (Kimura and Maruyama 1966); i.e., the number of individuals in the class with $m$ mutations is

$$n_m = N \exp \left( -u/s \right) \left( \frac{u}{s} \right)^m.$$  \hspace{1cm}

Assuming that the sequence of events is mutation, reproduction, and selection, populations were then run for 100 generations. After this initial period, populations were run for >2000 generations and up to 100,000 generations for conditions under which the ratchet clicks slowly, so that the average time between clicks of the ratchet could be measured. Every generation, the number of mutations in each individual is counted and the number of individuals with the least number of mutations (least-loaded class) is recorded. If, at a given generation, the number of mutations in the least-loaded class increases, the ratchet has clicked. To form a new generation, individuals are sampled randomly from the previous generation, then subjected to the occurrence of mutations sampled from a Poisson distribution with mean $u$, and assigned probabilities of survival as $(1 - s)^k$, where $k$ is the number of mutations that an individual carries. A new generation of $N$ individuals is constructed by comparing the probability of survival of each individual with a pseudorandom number drawn from a uniform distribution in the interval $[0, 1]$. Each run was repeated several times; generally five replicates were performed to obtain the results presented in the next section.

Although this simulation procedure does not follow the fate of each mutation at a particular locus, which is extremely time consuming, it gives the same results as the multilocus stochastic simulations of Charlesworth and Charlesworth (1997) as far as the estimation of the time between clicks of the ratchet is concerned, for all parameter sets tested (results not shown).

SIMULATION RESULTS

If the loss of the least-loaded class is the determining factor in driving the ratchet (Haigh 1978), one would expect that the time between clicks of the ratchet would stay approximately constant over a range of parameter values that keep $n_0$ constant. Figures 1, 2, and 3 show the simulation results for several parameter sets chosen such that $n_0$ stays constant.

In Figure 1, $s$ is kept constant at 0.015, and $N$ changes with $u$ to keep $n_0$ constant (either 20.2 or 202). We observe that the time between clicks of the ratchet does not change significantly over an order of magnitude change in $N$. The increase in $N$ seems to be compensated by the increase in $u$. In Figure 2, $u$ is kept constant, at 0.1, and $N$ changes together with $s$ to keep $n_0$ constant (with the same values as in Figure 1). Although for small
values of \( n_0 \) there seems to be no significant difference in the speed of the ratchet over an order of magnitude change in population size, for a higher value of \( n_0 \) the difference is evident: as \( s \) becomes large (small \( N \) in the plot), the speed of the ratchet is greatly reduced. For example, for a population size of 705 individuals, with a selection coefficient of 0.08, we did not observe any click over 50,000 generations, but for a population of 19,000 individuals, with \( s = 0.022 \), the average time for a click is \( \sim 1560 \) generations. In Figure 3, \( N \) is kept constant and \( u \) changes with \( s \) (the mean equilibrium number of mutations, \( u/s \), has the value 5), so that \( n_0 \) is constant. We see that, for either small \( u \) and small \( s \) or for large \( u \) and large \( s \), the speed of the ratchet is greatly reduced.

These results show that, as noted previously by STEPHAN et al. (1993), the size of the best class is not sufficient to predict the speed of the ratchet, because the ratchet can turn at very different speeds for a constant \( n_0 \). One observes that for the same \( n_0 \), keeping \( N \) constant and varying \( u \) and \( s \) so that \( u/s \) is constant, there is a value of \( u \) and \( s \) for which the time of the ratchet has a minimum, i.e., for the same \( n_0 \), increasing \( s \) can both slow down and speed up the ratchet (see the U-shaped curves in Figure 3). This shape is not predicted by any of the previous formulas. In contrast, such a minimum is predicted by the approximation presented here, although in the region of very small \( u \) and very small \( s \) the approximation gives lower times than the simulations. It also underestimates the time between clicks of the ratchet for small population size (or small \( n_0 \)).

The possible reasons why this minimum is observed, and why the present approximation underestimates the time for small population size, deserve some comment. Other things being equal, decreasing \( s \) should speed up the ratchet and decreasing \( u \) should slow it down. In the case of Figure 3, both \( s \) and \( u \) are changing to keep \( u/s \) constant, so that a minimum may occur, due to the fact that the dependence of the time on the mutation rate is different from that on the selection coefficient. In the region where \( s \) is very small, so that each mutation has very little effect on fitness, \( u \) is also very small. This means that the probability of a mutation occurring is very small, and the force of mutation that drives individuals from the best class to the next class is greatly reduced, leading to a slower ratchet. In the region where the mutation rate is large, the selection coefficient is also large, so that although mutations keep appearing at a high rate, selection is so efficient in restoring the best class that a great number of individuals come from the least-loaded class, which leads to a slower ratchet.

One observes from the comparison of the theoretical formula and the simulation results that, as long as \( s \) is not extremely small or large (Figures 1–3), the predictions seem to approximate the simulations reasonably well, especially when \( N \) is big (or \( n_0 \) is large). If \( N \) is small, it is more difficult for the system to maintain itself close to equilibrium, because drift is dominating. In this case, an approximation based on small perturbations becomes inadequate.

We now ask how the speed of the ratchet changes with population size, for a given mutation rate and a constant selection coefficient. The simulation results and the expected times calculated with the various ap-
propositions discussed above are shown in Figures 4–6, for different values of \( N, u, \) and \( s, \) as a function of \( n_0. \) The mutation rate and selection coefficient were chosen to lie in the parameter range that may be most relevant to the problem of the evolutionary degeneration of an incipient \( Y \) chromosome or neo-\( Y \) chromosome.

One observes that the approximation of Charlesworth and Charlesworth, contrary to what was previously thought (Charlesworth and Charlesworth 1997; Orr and Yuseob 1998), greatly underestimates the speed of the ratchet for large population sizes in the parameter range considered here. Although Stephan et al. could not establish exactly the range of validity of their two approximations, they suggested use of their Equation 8 for predicting the speed of the ratchet for the range of selection coefficients considered here. From comparison with the simulations presented, we see that their Equation 14 seems to describe the rate of change with \( N \) of the time between clicks of the ratchet better than their Equation 8, although it always underestimates the absolute time for this parameter range.

In simulations done to check the change in the ratchet’s speed with different mutation rates (Figure 7), we can see that the range of parameters for which Stephan et al.’s Equation 14 gives a better quantification of the process than their Equation 8 is not only dependent on a large population size and strong selection, but is also
modulated by the mutation rate. In simulations done to check the change of the ratchet's speed with different \( s \) (Figure 8) we see that, for a given \( N \) and \( u \), as \( s \) increases, the estimated time given by their Equation 14 becomes an underestimate. In fact, Stephan et al. studied the process by dividing it in two separate phases: the establishment phase, during which the new least-loaded class reaches a value close to that of the deterministic equilibrium, and the extinction phase, during which the least-loaded class becomes extinct. Their Equation 14 is based on the assumption that the population size is large enough that the change in mean fitness of the population is sufficiently small to be approximated by small perturbations and that selection is strong enough that the process is mostly trapped in the extinction phase. This assumption is not supported by our simulations (see Figure 9). The time spent in the establishment phase is the one above the dashed line; the time spent in the extinction phase is the one below this line.

When deriving our approximation, we assumed that the net loss of mean fitness due to a click of the ratchet would be \( K = 0.6s \), because the system did not recover its new equilibrium instantaneously. In Figure 10A, we plot the mean fitness of the population as a function of \( 1 - x/x_0 \) for a set of simulation runs with parameters \( N = 10,000, u = 0.04, s = 0.01 \), after a click of the ratchet. As we assumed that the mean fitness of the population could be approximated as a linear function of \( 1 - x/x_0 \), the slope of the linear regression line plotted in the figure corresponds to the value of \( K \) in the theoretical approximation. The value assumed in the derivation (\( K = 5.76 \times 10^{-5} \)) agrees well with the one from the regression. However, the agreement is not good for a population size of only 1000 (Figure 10B). This can be attributed at least to two factors: either the population size is so small that we cannot approximate the changes in mean fitness by small perturbations and/or the value of \( K \) is different from the one we are assuming. If we calculate the time by substituting the value of \( K \) from the linear fit in the theoretical expression, we find that the time obtained is still below the one measured in the simulations, so that an incorrect value for \( K \) is not the only source of the discrepancy.

Figure 7.—The time between clicks of the ratchet, ±2 SE, as a function of the mutation rate (\( N = 10,000, s = 0.015 \)).

Figure 8.—The time between clicks of the ratchet, ±2 SE, as a function of the selection coefficient (\( N = 10,000, u = 0.075 \)).

Figure 9.—The fluctuations in the size of the least-loaded class through time, for \( N = 10,000, u = 0.075 \), and \( s = 0.02 \), for a single simulation run. The dashed line indicates the equilibrium value of the least-loaded class.
discretion to ensure 
that the Poisson distribution expected under mutation-selection balance can be ap-
proximately attained. GESSLER (1995) has shown that 
this balance will not be met for conditions under which 
\( n_b < 1 \), because selection is too weak to counter mutation 
pressure. In this case, the distribution of the number of mutations is not Poisson but is close to a shifted 
negative binomial distribution, whose parametrization allows an estimation of the rate of the ratchet. 

For \( n_b \gg 1 \), the approximation for the advance of the ratchet based on a diffusion equation presented here 
seems to make a better prediction of the time between 


clicks of the ratchet than the previous approximations, 
for moderate selection coefficients in a range that is 
compatible with the biological data, provided the popu-
lation size is not too small, so that the stochastic fluctua-
tions are not too violent. As noted before by 
STEPHAN et al. (1993), for intermediate selection coefficients the 
establishment phase and the extinction phase are 
blurred and a separate analysis of these two phases does 
not well predict the outcome. An approximation based on 
the assumption that the mean fitness of the population 
is affected solely by fluctuations of the least-loaded class 
(\( w = (x - x_0) + e^{\omega} \)) (CHARLESWORTH and 
CHARLESWORTH 1997) seems to approximate the simulations 
reasonably well for a small equilibrium size of the best 
class (although this is highly dependent on the mutation 
rate), but greatly overestimates the time for a click when 
\( n_b \) is large or \( u \) is low. 

The observation that, for a constant \( n_b \) and constant 
\( s \), we did not observe significant differences in the speed 
of the ratchet, over an order of magnitude change in the 
population size (Figure 1), suggests that \( n_b s \) is an 
important parameter, although the expression for the 
average time between clicks of the ratchet is not an 
explicit function of \( n_b s \). For the parameter range consid-
ered here we observe that an increase of 10-fold in 
\( n_b s \) caused a decrease of \( \sim 10 \)-fold in the speed of the ratchet. 

\( Y \) and neo-\( Y \) chromosome degeneration: Because an 
inipient \( Y \) chromosome, or a neo-\( Y \) chromosome re-
sulting from an autosome fusion or translocation, that 
fails to recombine with its homologue in the heteroga-
metic sex is vulnerable to the ratchet, it is interesting 
to calculate the expected rate of its operation under 
the above approximation. The erosion of a proto-\( Y \) 
chromosome is very similar to the degeneration of a haploid 
asexual population if one replaces \( s \) by \( h s \), where \( h \) is the 
dominance coefficient and \( s \) is the selection coefficient 
against homozygous mutations (CHARLESWORTH and 
CHARLESWORTH 1997). In the case of Drosophila, if one 
assumes an effective population size of males of \( 5 \times 10^2 \), a deleterious mutation rate per \( Y \) chromosome of 
0.04, and an average selection coefficient against a het-
erozygous mutation of 0.01, the present approximation gives a value of \( \sim 3 \times 10^{10} \) generations for one click 
of the ratchet; if one sets 5 generations per year for 
Drosophila, this will correspond to \( 6 \times 10^4 \) years per click. However, if the mutation rate is slightly higher,
say 0.06, this time would decay to 40,000 years per click. Our approximation suggests that an increase in the mean number of mutations from 4 (n_b = 9158) to 6 (n_b = 1239), just due to an increase in the mutation rate, increases the speed of the ratchet by 20 orders of magnitude. This comes from the fact that there is a more or less exponential increase of the time between clicks of the ratchet with an increase in n_b for a given selection coefficient. On the other hand, for a mutation rate of 0.04, if hs is ~0.005, n_b will be 168 and the estimated average time for a click is 697 generations. Because every click of the ratchet in a haploid population indirectly leads to the fixation of a deleterious mutation in the whole population (Charlesworth and Charlesworth 1997), this means that if the ratchet is clicking every 697 generations, in a period of 1 million years we would expect an incipient Drosophila Y chromosome to become fixed for ~7170 deleterious mutations (again assuming 5 generations per year).

The neo-Y chromosome system of Drosophila miranda constitutes an excellent clock to set the time scale over which the degeneration of a nonrecombining region is supposed to occur. The time estimated for the origin of the chromosomal rearrangement generating the neo-Y in D. miranda is ~1.25 million years ago (mya; S. Yi, personal communication). The neo-Y shows evidence for degeneration, and the neo-X is partially dosage compensated (Steinemann et al. 1993; Steinemann and Steinemann 1998, 1999). These observations seem to suggest that, if there is a general process responsible for the degeneration of the nonrecombining segment of the genome such as Muller's ratchet, it is expected to show its signature over a time scale of the order of 10^6 years.

In Table 1 we show the expected time for a click of the ratchet, under the approximation proposed here, for a population of half a million males for various values of u and hs. We observe that, for these values for the average effect of a mutation, the time for a click of the ratchet becomes biologically irrelevant when n_bhs goes above 15. If the average heterozygous effect of a nonlethal deleterious mutation is of the order of 1% (Charlesworth and Hughes 1999), under the above approximation, for n_b < 1500 the ratchet may play a role in the degeneration of the neo-Y, but its rate is probably too small to explain the degeneration observed if n_b > 1000. For the ratchet to be the main process causing the degeneration of the neo-Y, n_b has to be probably <500. In the case of lethal mutations, which probably occur at a rate of ~0.0025 for an incipient Y chromosome (Fry et al. 1999) and with hs of ~2% (Crow 1993), the time for a click of the ratchet in a population of half a million chromosomes is biologically irrelevant.

Muller's ratchet is a priori more likely to be an important force in driving the degeneration of mammalian Y chromosomes, given that their effective population size is around one order of magnitude less than that for Drosophila (Charlesworth and Charlesworth 1997). There is some evidence that the evolution of the mammalian Y chromosome has been punctuated by at least four events that suppressed recombination between the X and the Y (Lahn and Page 1999), the first event having occurred ~300 mya. If the deleterious mutation rate and average effect of mutations for a mammalian proto-Y chromosome were the same as those estimated for the Y in Drosophila, then with a population of 5 \times 10^4, proto-Y chromosomes would degenerate due to this process at an average rate of 1 click every 40 thousand generations, for a deleterious mutation rate of 0.04. Our ignorance of the value of these parameters does not allow any final conclusion about the ratchet being a leading process in the degeneration of the mammalian Y, although it seems more likely than for the Drosophila case.

Because of the assumptions we have made to derive these results, some caution has to be taken considering their implications. First we must note that we have assumed, for simplicity, that all deleterious mutations have the same effect. However, recent work has suggested that an equal effect of mutations assumption does not fit the data from mutation accumulation experiments, which are designed to measure the mutation rate to deleterious mutations and the selection coefficients against those mutations (Keightley 1996; Davies et al. 1999; Fry et al. 1999). The occurrence of many mutations with small effects, and a few with large effects, seems to be more consistent with the results. If this is the case, the ratchet is expected to turn at a much higher speed than for a single selection coefficient of the order of 1% (Gessler and Xu 1999), but each turn

<table>
<thead>
<tr>
<th>u</th>
<th>hs</th>
<th>n_b</th>
<th>Time</th>
<th>n_bhs</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.03</td>
<td>0.005</td>
<td>1239</td>
<td>2 \times 10^4</td>
<td>6</td>
</tr>
<tr>
<td>0.04</td>
<td>0.005</td>
<td>168</td>
<td>697</td>
<td>0.84</td>
</tr>
<tr>
<td>0.04</td>
<td>0.008</td>
<td>3369</td>
<td>1 \times 10^6</td>
<td>27</td>
</tr>
<tr>
<td>0.04</td>
<td>0.01</td>
<td>9158</td>
<td>3 \times 10^9</td>
<td>92</td>
</tr>
<tr>
<td>0.05</td>
<td>0.01</td>
<td>3369</td>
<td>4 \times 10^18</td>
<td>34</td>
</tr>
<tr>
<td>0.06</td>
<td>0.01</td>
<td>1239</td>
<td>2 \times 10^18</td>
<td>12</td>
</tr>
<tr>
<td>0.07</td>
<td>0.01</td>
<td>456</td>
<td>4 \times 10^18</td>
<td>5</td>
</tr>
<tr>
<td>0.07</td>
<td>0.015</td>
<td>4702</td>
<td>7 \times 10^18</td>
<td>71</td>
</tr>
<tr>
<td>0.08</td>
<td>0.015</td>
<td>2414</td>
<td>1 \times 10^21</td>
<td>36</td>
</tr>
<tr>
<td>0.09</td>
<td>0.015</td>
<td>1239</td>
<td>5 \times 10^20</td>
<td>19</td>
</tr>
<tr>
<td>0.1</td>
<td>0.015</td>
<td>636</td>
<td>3 \times 10^19</td>
<td>10</td>
</tr>
<tr>
<td>0.1</td>
<td>0.02</td>
<td>3369</td>
<td>9 \times 10^19</td>
<td>67</td>
</tr>
<tr>
<td>0.13</td>
<td>0.02</td>
<td>752</td>
<td>5 \times 10^19</td>
<td>15</td>
</tr>
</tbody>
</table>

Average time for a click of the ratchet (in generations), predicted by the proposed approximation, for a population of 500,000 neo-Y chromosomes with various mutation rates and selection coefficients.
will cause a very small decline in mean fitness if many mutations have very low selection coefficients. It is likely that the degeneration of the Y chromosome and the evolution of dosage compensation are driven by selection to increase the activity of the X relative to the Y in males, in response to the decline in mean fitness of the Y (Charlesworth 1996; Charlesworth and Charlesworth 1997). The strength of such selection is determined by the rate of this decline and will be very weak if it is small.

We have also assumed independence of mutational effects and it has been shown that epistasis slows down the speed of this process (Charlesworth et al. 1993; Kondrashov 1994). However, if there is in fact a distribution of mutational effects with a more or less exponential shape, epistasis will not stop the ratchet (Butcher 1995).

We thank D. Gessler and E. Baake for helpful comments. I. Gordo also thanks C. McVean for helpful discussions concerning this work. I. Gordo is supported by the Gulbenkian Foundation and Program PRAXIS XXI of Portugal, and B. Charlesworth by the Royal Society.

LITERATURE CITED


Charlesworth, B., and D. Charlesworth, 1997 Rapid fixation of deleterious alleles can be caused by Muller’s ratchet. Genet. Res. 70: 63–73.


Communicating editor: W. Stephan
Letter to the Editor

On the Speed of Muller's Ratchet

Isabel Gordo and Brian Charlesworth

Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom

Manuscript received May 29, 2000
Accepted for publication July 26, 2000

WILE asexual populations can suffer from an effectively irreversible accumulation of mildly deleterious mutations, sexual populations are essentially immune to it. This remarkable difference between the absence and the presence of recombination was first put into words by Muller, who defined what later was named Muller's ratchet (Muller 1964; Felsenstein 1974). Since then, Muller's ratchet has been proposed as a potential explanation for the advantage of recombination, the extinction of asexual populations, a limit to the genome size of asexual organisms, and the degeneration of Y chromosomes (Charlesworth 1978; Maynard Smith 1978; Lynch et al. 1993; Charlesworth and Charlesworth 1997; Gessler and Xu 1999). In its classical formulation, the ratchet is portrayed as a process of successive losses of the least-loaded classes of individuals (the class with the minimum number of mutations at any one time) due to genetic drift. One important question concerns the speed with which it operates, usually defined as the mean time for one turn of the ratchet. This was first examined by Haigh (1978), who identified the following relevant parameters for its quantification: the population size (N), the mutation rate (u), the selective effect of a single mutation (s), and the size of the least-loaded class at mutation-selection equilibrium n0 = Nexp(−u/s), assuming multiplicative fitnesses.

Although a general expression for the speed of this process remains to be obtained, we have recently provided an expression for quantifying the mean time between turns of the ratchet that appeared to be a good approximation for moderate values of s and for n0 ≈ 1 (Gordo and Charlesworth 2000). Since it has recently been proposed that there may be a large class of mildly deleterious mutations with selection coefficients <2% (Keightley and Eyre-Walker 1999), we need to have an approximation to the speed of the process for this parameter range. Here we reexamine our previous approximation and suggest a more robust prediction that seems to work better over a wider range of parameters. As in previous investigations (Stephan et al. 1993; Charlesworth and Charlesworth 1997), we have tried to determine the speed of Muller's ratchet by modeling the ratchet as a one-dimensional diffusion process for which we calculate the mean time to absorption of the frequency of the least-loaded class (Gordo and Charlesworth 2000). To derive the diffusion coefficient, we assume a Wright-Fisher model; for the drift coefficient, we assume small perturbations around the equilibrium under deterministic mutation-selection balance. For the latter, we also assume that, just after a turn of the ratchet, the size of the new least-loaded class would rapidly approach a value close to 1.6 n0 as suggested by the corollary to Haigh's (1978) Theorem 1. When s is intermediate, this is likely to be true, but the smaller the value of s, the longer the time it takes to approach this value. In particular, when n0 and s are small, then this relaxation time becomes the main determinant of the total time to absorption (Haigh 1978; Stephan et al. 1993). It is this additional time (T0), which we previously assumed to be effectively zero, that we must add to get a better prediction.

We can approximate T0 by the time it takes to get from the size of the new least-loaded class immediately after one turn (which at this point has an approximate value of n1 = n0 u/s) to 1.6 n0 using Haigh's Theorem 1 or Equation 3 in Stephan et al. (1993). This is approximately

\[
T_0 = \frac{1}{s} \left( 1 - \frac{1.6}{u} \right) \tag{1}
\]

Therefore, the mean time for a turn of the ratchet is T(N, u, s) = T0 + T0.01 + T0.1, where T0.01 is the time spent in the frequency interval [0, x0] and T0.1 is the...
Comparison of the mean time between turns of the ratchet between simulations, $T$ (with 2 SE), and the analytical expression, $T(N, u, s)$

<table>
<thead>
<tr>
<th>$N$</th>
<th>$u$</th>
<th>$s$</th>
<th>$n_0$</th>
<th>$T$ (2 SE)</th>
<th>$T(N, u, s)$</th>
<th>$\omega/\omega^*$</th>
<th>No. of fixations$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>275</td>
<td>0.06</td>
<td>0.015</td>
<td>5</td>
<td>63 (5)</td>
<td>66</td>
<td>$8.06 \times 10^{-5}$</td>
<td>7,937</td>
</tr>
<tr>
<td>565</td>
<td>0.1</td>
<td>0.03</td>
<td>20</td>
<td>119 (2)</td>
<td>101</td>
<td>$3.29 \times 10^{-5}$</td>
<td>4,119</td>
</tr>
<tr>
<td>1,000</td>
<td>0.04</td>
<td>0.01</td>
<td>18</td>
<td>169 (39)</td>
<td>144</td>
<td>$1.22 \times 10^{-1}$</td>
<td>2,959</td>
</tr>
<tr>
<td>3,000</td>
<td>0.0075</td>
<td>0.0015</td>
<td>20</td>
<td>489 (42)</td>
<td>571</td>
<td>$2.15 \times 10^{-1}$</td>
<td>1,023</td>
</tr>
<tr>
<td>5,000</td>
<td>0.1</td>
<td>0.02</td>
<td>20</td>
<td>135 (14)</td>
<td>120</td>
<td>$3.19 \times 10^{-3}$</td>
<td>3,704</td>
</tr>
<tr>
<td>6,065</td>
<td>0.03</td>
<td>0.005</td>
<td>15</td>
<td>215 (25)</td>
<td>225</td>
<td>$8.66 \times 10^{-6}$</td>
<td>2,326</td>
</tr>
<tr>
<td>10,000</td>
<td>0.075</td>
<td>0.01</td>
<td>5.5</td>
<td>126 (12)</td>
<td>109</td>
<td>$4.78 \times 10^{-18}$</td>
<td>3,968</td>
</tr>
<tr>
<td>13,500</td>
<td>0.075</td>
<td>0.02</td>
<td>317</td>
<td>4,574 (1,913)</td>
<td>4,712</td>
<td>$1.10 \times 10^{-1}$</td>
<td>109</td>
</tr>
<tr>
<td>30,000</td>
<td>0.005</td>
<td>0.001</td>
<td>202</td>
<td>1,743 (225)</td>
<td>1,601</td>
<td>$6.98 \times 10^{-1}$</td>
<td>359</td>
</tr>
<tr>
<td>30,000</td>
<td>0.01</td>
<td>0.002</td>
<td>202</td>
<td>1,225 (272)</td>
<td>1,200</td>
<td>$4.42 \times 10^{-3}$</td>
<td>408</td>
</tr>
<tr>
<td>30,000</td>
<td>0.1</td>
<td>0.02</td>
<td>202</td>
<td>1,543 (275)</td>
<td>1,521</td>
<td>$1.44 \times 10^{-3}$</td>
<td>324</td>
</tr>
<tr>
<td>30,000</td>
<td>0.175</td>
<td>0.035</td>
<td>202</td>
<td>4,890 (454)</td>
<td>3,860</td>
<td>$2.62 \times 10^{-2}$</td>
<td>102</td>
</tr>
<tr>
<td>50,000</td>
<td>0.04</td>
<td>0.01</td>
<td>918</td>
<td>45,922 (26,244)</td>
<td>38,957</td>
<td>$8.96 \times 10^{-1}$</td>
<td>11</td>
</tr>
<tr>
<td>500,000</td>
<td>0.015</td>
<td>0.0025</td>
<td>1,259</td>
<td>4,924 (764)</td>
<td>7,785</td>
<td>$7.76 \times 10^{-1}$</td>
<td>102</td>
</tr>
<tr>
<td>500,000</td>
<td>0.015</td>
<td>0.0015</td>
<td>23</td>
<td>732 (66)</td>
<td>1,343</td>
<td>$3.59 \times 10^{-1}$</td>
<td>683</td>
</tr>
<tr>
<td>500,000</td>
<td>0.03</td>
<td>0.003</td>
<td>23</td>
<td>490 (114)</td>
<td>716</td>
<td>$4.66 \times 10^{-2}$</td>
<td>1,020</td>
</tr>
<tr>
<td>500,000</td>
<td>0.3</td>
<td>0.03</td>
<td>23</td>
<td>167 (79)</td>
<td>139</td>
<td>$4.29 \times 10^{-3}$</td>
<td>2,994</td>
</tr>
<tr>
<td>500,000</td>
<td>0.03</td>
<td>0.005</td>
<td>1,259</td>
<td>12,819 (5,681)</td>
<td>17,625</td>
<td>$8.22 \times 10^{-1}$</td>
<td>39</td>
</tr>
<tr>
<td>500,000</td>
<td>0.04</td>
<td>0.004</td>
<td>23</td>
<td>371 (86)</td>
<td>570</td>
<td>$4.51 \times 10^{-3}$</td>
<td>1,348</td>
</tr>
<tr>
<td>500,000</td>
<td>0.04</td>
<td>0.005</td>
<td>168</td>
<td>940 (106)</td>
<td>1,018</td>
<td>$6.95 \times 10^{-1}$</td>
<td>532</td>
</tr>
<tr>
<td>500,000</td>
<td>0.04</td>
<td>0.006</td>
<td>656</td>
<td>3,688 (322)</td>
<td>4,657</td>
<td>$4.59 \times 10^{-1}$</td>
<td>129</td>
</tr>
<tr>
<td>500,000</td>
<td>0.04</td>
<td>0.0065</td>
<td>1,063</td>
<td>15,381 (6,210)</td>
<td>19,225</td>
<td>$7.84 \times 10^{-1}$</td>
<td>37</td>
</tr>
<tr>
<td>500,000</td>
<td>0.05</td>
<td>0.0075</td>
<td>636</td>
<td>3,912 (369)</td>
<td>5,910</td>
<td>$3.80 \times 10^{-1}$</td>
<td>128</td>
</tr>
<tr>
<td>500,000</td>
<td>0.056</td>
<td>0.008</td>
<td>456</td>
<td>2,420 (303)</td>
<td>3,219</td>
<td>$1.90 \times 10^{-1}$</td>
<td>207</td>
</tr>
<tr>
<td>500,000</td>
<td>0.065</td>
<td>0.01</td>
<td>752</td>
<td>9,310 (1,306)</td>
<td>16,934</td>
<td>$5.83 \times 10^{-1}$</td>
<td>54</td>
</tr>
<tr>
<td>500,000</td>
<td>0.07</td>
<td>0.01</td>
<td>456</td>
<td>2,711 (411)</td>
<td>4,022</td>
<td>$1.57 \times 10^{-1}$</td>
<td>184</td>
</tr>
<tr>
<td>500,000</td>
<td>0.08</td>
<td>0.01</td>
<td>168</td>
<td>812 (113)</td>
<td>915</td>
<td>$2.05 \times 10^{-3}$</td>
<td>616</td>
</tr>
</tbody>
</table>

$^a$The expected ratio of the mean fitness of the population and the initial mean fitness after 500,000 generations.

$^b$Expected number of fixations of deleterious mutations after 500,000 generations.

time spent in the interval $[n_0, 1]$, given by Equations 3a and 3b in GORDO and CHARLESWORTH (2000).

While $T_0$ is the deterministic time for the frequency to approach a state close to the new mutation-selection balance, the other terms represent the mean time of the stochastic process leading to absorption. For a given $N$ and $u$, small values of $n_0$ correspond to small values of $s$, and $T_0$ dominates the other terms; as $s$ increases, so does $n_0$, and the value of $T_0$ becomes less relevant compared with the other terms.

In Table 1 we compare the results of this formulation with those obtained by simulations. The simulation method is as described in GORDO and CHARLESWORTH (2000), i.e., we assume the sequence of events: mutation, reproduction and selection, a constant population size, and multiplicative fitness effects of mutations at different loci. The parameter values were chosen as follows. The large values of $N$ in Table 1 are of considerable biological importance, since we want to analyze the role of Muller's ratchet in the degeneration of Y and neo-Y chromosomes in systems such as *Drosophila miranda*, for which the effective population size is thought to be in the order of hundreds of thousands or millions (Y1 and CHARLESWORTH 2000). A Y or neo-Y chromosome behaves like a haploid asexual population, for which $N$ is the number of breeding males, the selection coefficient is the effect of a mutation when heterozygous, and the mutation rate is the fraction of the total haploid deleterious mutation rate corresponding to the size of that chromosome (CHARLESWORTH 1978). The values of the deleterious mutation rate were chosen to cover a region that is reasonable in the light of various mutation-accumulation experiments (KEIGHTLEY and EYRE-WALKER 1999). All the parameters are constrained to the condition $n_0 > 1$, because, if this condition is not met, the assumption of the existence of a state characterized by the deterministic mutation-selection balance is invalid (see GESSLER (1995) for results on the case of $n_0 < 1$). For all the simulation results reported previously (GORDO and CHARLESWORTH 2000) and other parameters that we tested, the new expression performs better.

In Figure 1 we show, as an example, the dynamics of the size of the least-loaded class over time intervals of
10 generations after a turn of the ratchet, taken from several simulation runs. The parameter values are \( N = 10,000 \), \( u = 0.03 \), and \( s = 0.005 \). With these parameters, \( n_0 = 25 \), \( 1.6 n_0 = 40 \), and \( n_1 = 149 \). Although there is an enormous variance in the behavior of the changes in size of the least-loaded class, on average (thick line in the figure) the behavior is close to what we have assumed. Immediately after a turn of the ratchet, the mean size of the least-loaded class is close to \( n_0 \), and then it approaches a value close to \( n_0 \) over 100–200 generations. This pattern is essentially the same for other parameter values. One fact is probably worth noting: although we can, with a single expression, estimate reasonably well the time between turns of the ratchet (for very different values of \( N \), \( u \), and \( s \)), when \( s \) is large (>0.04) our expression underestimates the time obtained in the simulations. In this range, none of the diffusion approximations is accurate, as expected from conditions for diffusion theory to be reliable (Ewens 1979, Chap. 4).

Under this model, under which each mutation causes an identical and independent deleterious effect on fitness, the decline in the logarithm of mean fitness is

\[
\frac{\partial \ln \bar{w}}{\partial t} = \frac{\ln(1 - s)}{T} = \frac{s}{T}
\]

where \( T \) is the mean time for a turn of the ratchet.

Clearly, deleterious mutations with larger effects cause a bigger decline in log mean fitness per turn but take more time to accumulate, while weaker deleterious mutations will accumulate faster but cause a smaller decline in log mean fitness (as noted before by Lynch et al. 1993). Therefore there is a value of \( s \), say \( s_{\text{max}} \), that maximizes the decline in log mean fitness. The partial derivation of \( s/T \) with respect to \( s \) will be zero at this point. We can easily calculate the approximate value of \( s_{\text{max}} \), using our approximation. This value is obviously a function of \( N \) and \( u \). For example, in the case of a population of Drosophila \( Y \) or neo-\( Y \) chromosome, \( N \) is likely to be of the order of 500,000 if we assume an effective population size for Drosophila of \( \sim 1 \) million and a 1:1 sex ratio (\( Y \) and Charlesworth 2000). If the mutation rate is 0.04, then mutations whose effect is \( \sim 0.004 \) are expected to cause the biggest decline in log mean fitness \( (s/T = 3.7 \times 10^{-5}) \). If \( u \) is smaller, say 0.02, then weaker mutations will correspond to the maximum rate of decline but cause a much lower rate of decline \( (s/T = 3.7 \times 10^{-5}) \) than in the first case. From (2), we can calculate the ratio of the mean fitness at any time to the initial mean fitness of the population, \( w/w_0 \). We display this ratio after 500,000 generations in Table 1. We also show the expected number of fixed deleterious mutations at this time, since it is known that, in the long run, the rate of the ratchet is the rate of fixation of deleterious mutations (Charlesworth and Charlesworth 1997).

For large populations, the average time between turns of the ratchet, for mutations that cause a considerable decline in log mean fitness of the population per turn \( (0.005 < s < 0.01) \), is on the order of thousands of generations (see Table 1) for values of \( u \) that are possibly reasonable for large nonrecombining segments of the genome (such as the \( Y \) chromosome) in real populations of this size. The neo-\( Y \) chromosome of \( D. miranda \) results from a fusion between an autosome and the \( Y \) chromosome, and the estimated time of origin of the rearrangement is \( \sim 1 \) million years ago (\( Y \) and Charlesworth 2000). Since there are \( \sim 13,600 \) genes in Drosophila (Rubin et al. 2000) and the neo-\( Y \) constitutes about one-fifth of the genome, we expect \( \sim 2700 \) genes on the neo-\( Y \). This means that, if the ratchet is operating approximately as in our model, we expect hundreds of fixations of mildly deleterious mutations in about one-tenth of the total lifetime of the neo-\( Y \). Contrary to the suggestion of Charlesworth (1996), the ratchet thus seems to be a viable mechanism for the degeneration of the neo-\( Y \) if the great majority of deleterious mutations have selection coefficients <2%. Of course this process is not incompatible with the operation of other processes (as discussed in Charlesworth 1996).

**LITERATURE CITED**


Charlesworth, B., and D. Charlesworth, 1997 Rapid fixation.
of deleterious alleles can be caused by Muller's ratchet. Genet. Res. 70: 63–73.


Communicating editor: N. Takahata
Primer

Genetic linkage and molecular evolution
Isabel Gordo and Brian Charlesworth

The rate of genetic recombination varies among species, and across different regions of the genome within a species. In the extreme case, all genes in an asexual or self-fertilizing species are effectively completely linked, in contrast to the substantial opportunities for recombination in outbreeding sexual species. Several key observations on the relation between patterns of evolution and levels of genetic recombination have been made. First, asexual or highly self-fertilizing species tend to be young in terms of the evolutionary timescale, suggesting that they become extinct more rapidly than their recombining relatives. Second, organelle and clonally transmitted genomes in several taxa show reduced levels of adaptation with respect to RNA and protein sequences. In Drosophila, regions where the recombination rate is low, such as near centromeres and telomeres, show reduced levels of codon bias — the non-random use of alternative codons encoding the same amino acid — suggesting a reduction in the ability of selection to maintain this aspect of molecular adaptation. Finally, in a number of different species, genomic regions with restricted recombination show lower levels of DNA sequence variation. One of the best examples is provided by non-recombining Y chromosomes, or neo-Y chromosomes formed by fusions between autosomes and sex chromosomes, which show degeneration of gene function and reduced genetic variability.

Why should there be a relation between the amount of recombination and levels of variation and adaptation? As Fisher and Muller pointed out, the dynamics of a given gene are influenced by the evolutionary forces acting on the gene itself, as well as by forces acting at linked loci. Thus the predictions of single-locus population genetics must be modified when selection is acting on sets of linked loci. Here we describe some of the processes that may shape evolution when recombination is restricted over a large genomic region, and which may explain the above observations. These processes all reflect a general effect first quantified by Hill and Robertson: a locus linked to another locus under directional selection experiences a reduced effective population size, $N_e$. The extent of random fluctuations in allele frequencies due to finite population size, genetic drift, is inversely related to $N_e$; thus selective differences at one locus tend to enhance the effects of drift at a linked locus.

According to the neutral theory of molecular evolution, mutation and genetic drift have a major influence on DNA sequence variation within species and on differences between species. Mutation creates new neutral variants, with no significant fitness effects, and genetic drift causes random changes in their frequencies until fixation or loss. Kimura showed that the level of variation within a population at a neutral locus is proportional to the product of $N_e$ and the neutral mutation rate, $\mu$, and that the rate of sequence evolution is equal to $\mu$. A mutation with a selection coefficient, $s$, which measures the reduction or increase in the fitness of its carriers relative to that of the rest of the population, is effectively neutral if $N_e s < 1$. A strongly deleterious mutation, for which $N_e s >> 1$, will be rapidly eliminated, but a weakly deleterious mutation, with $N_e s < 1$, can persist and even become fixed in the population through genetic drift. Similarly a favourable mutation will have almost the same chance of loss from the population as a neutral mutation if $N_e s < 1$. Changes in $N_e$ caused by different kinds of selection at linked loci can thus greatly affect both genetic variability and the efficacy of selection.

Selective sweeps
Consider the effect of an advantageous mutation on the level of variation at a completely linked
neutral locus (Figure 1). A peculiar footprint is left just after such a mutation has swept through the population to fixation: variation is drastically reduced, and any variants that can be observed occur at low frequencies in the population. From the point of view of the neutral locus, it is as if the population went through a bottleneck of one individual, and then expanded to its normal size. If an advantageous mutation occurs in a region where there is a small amount of recombination, neutral variation will not be completely lost, but will still be reduced. The level of variation in a particular region of a chromosome depends on the rate of adaptive evolution and on the ratio of the strength of selection to the recombination rate in that region. If selective sweeps continually occur throughout the genome of a sexual species, there will be a correlation between neutral variability and the local recombination rate. Recurrent selective sweeps will also cause very low levels of variation in an asexual species, even if it has a large population size. Sweeps can also result in the fixation of linked deleterious alleles, and this has been proposed as a possible factor in the degeneration of Y chromosomes.

Background selection

Any natural population is subject to a continual rain of deleterious mutations at loci throughout the genome. If such mutations are strongly selected against, they will be reduced by a factor of \(N_e\). The rate of neutral evolution is independent of \(N_e\) and so remains unchanged. Now consider the effect of background selection when there is a certain frequency of recombination, \(r\), between a neutral variant and a deleterious mutation with which it is associated. The neutral variant can now unhitch itself from the deleterious mutation through recombination. As the mean time spent by a deleterious allele in a large population is roughly \(r^{-1}\), the neutral variant has a chance of survival if it unhitches itself during this time. Thus \(r/s\) is an important ratio in determining the effect of background selection. Mathematical analysis shows that this effect can be indeed approximated by a reduction in \(N_e\) that involves \(r/s\), and yields an expression relating the level of variation in a region of the genome to the local rate of recombination. This can largely explain the Drosophila data mentioned at the start.

If the sites concerned are not evolving neutrally but are very weakly selected, as with synonymous changes to codons, the effect of background selection can be quantified as though \(N_e\) is reduced in the same manner as for neutral sites. Because the rate of evolution at weakly selected sites depends on \(N_e\), the rate of fixation of weakly deleterious mutations is increased, and that of advantageous mutations is decreased, if background selection is operating. In regions of the genome where linkage is very tight, such as the Y chromosome, levels of variability and adaptation should both be reduced. Because background selection impairs the ability of an asexual species to adapt, it may also accelerate the extinction of asexual populations.
Muller's ratchet
The background selection model assumes that the frequency of the least-loaded class is stable. But this need not be the case. When \( f_s \) is small, Muller pointed out that genetic drift could then cause the loss of the least-loaded class — a 'click' of the ratchet. After this click, the class with one deleterious mutation becomes the new least-loaded class. But this class can also be lost in the same way as the previous one, so that successive clicks occur, leading to a continual accumulation of deleterious mutations to the detriment of the population's mean fitness — a process now known as Muller's ratchet. As a consequence of each click, fixation of a deleterious allele occurs in the entire population; such fixations would effectively be impossible in the absence of the ratchet. As with background selection, the ratchet can cause a strong reduction in the level of variability at neutral loci.

Muller argued that recombination is the only way to recreate the fittest individuals, and suggested that avoidance of the ratchet provides an advantage to recombination. The ratchet may also be involved in Y chromosome degeneration and the extinction of asexual populations. The frequency at which the ratchet clicks is a very important determinant of its evolutionary effects. Theoretical work shows that the ratchet clicks faster with smaller population size, higher mutation rates, and smaller selection coefficients. Very large asexual species with small genomes are probably immune to the ratchet, but species of smaller sizes with large genomes and no recombination will probably suffer from its consequences.

Weak Hill-Robertson effects
Suppose now that many sites in a tightly linked genome are undergoing mutations to very weakly selected alleles, for which \( N_r s \) is around one. Such mutations segregate at much higher frequencies than those involved in the previous processes, and there is now a non-negligible chance of back-mutation. An example of this is provided by synonymous mutations affecting codon usage, which may affect the efficiency or accuracy of translation. Models of genomes with many sites subject to such weakly selected mutations show that, as linkage increases, both intra-species variability and the mean level of adaptation — measured by the frequency with which optimal codons are used — decrease. This effect results from the cumulative effect of numerous polymorphic selected sites on \( N_r \), and is larger as more sites are involved. This process may be of importance in the evolution of any sizeable genome where linkage is complete, and in genomic regions with low levels of recombination.

The effect of balancing selection
There is one situation, however, where selection at a locus can increase the level of variability at linked neutral sites. This is when two or more alternative alleles at the selected locus are maintained in the population by balancing selection, for far longer than they would persist if they were neutral. Here, the different alleles act as separate sub-populations, which diverge under drift and mutation, leading to increased variability in the population as a whole. Recombination between a neutral locus and the selected locus is analogous to migration, and opposes this divergence. The size of the increase in variability is proportional to \((N_r r)^{-1}\). We thus expect increased variation in the neighbourhood of a target of such balancing selection, falling off with the genetic map distance from the target; this is seen in the MHC complex in humans, and self-incompatibility loci in plants.

Relating models to observations
The different models make similar predictions which makes it difficult to ascribe an observation to just one process. Are there any observations that allow one to identify which process is operating? It turns out that the distribution of frequencies of neutral variants in a sample of sequences (the frequency spectrum) is model-dependent. A large distortion of the frequency spectrum towards rare variants is more likely with selective sweeps (Figure 1). And when there is some recombination, so that initially rare variants are not necessarily swept to complete fixation, a transient signature of a sweep is provided by the occurrence of an excess of 'derived' variants at a high frequency within a sample. (One can infer if a variant is derived or ancestral from the sequence of a closely related species.) Other features of genetic variation, such as patterns of linkage disequilibrium, may also be useful in testing alternative models. Perhaps the best way to quantify the relative importance of these processes is to get solid estimates of the rate at which deleterious mutations occur, and of the distribution of their effects on fitness. Although this is a simple question to ask, it is hard to answer. But given such information, and using the increasing amount of information on DNA sequence variation and evolution, one can perhaps try to answer an even harder question: what is the rate at which advantageous mutations occur, and what are their effects on fitness?

Key references

Address: Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, UK.