Genetica **102/103:** 507–524, 1998.

Mutation and selection within the individual

Sarah P. Otto[1] & Ian M. Hastings[2]
[1] *Department of Zoology, University of British Columbia, 6270 University Blvd., Vancouver BC, Canada V6T 1Z4 (Phone: (604) 822-2778; Fax: (604) 822-2416; E-mail: otto@zoology.ubc.ca);* [2] *Institute of Cell, Animal, and Population Biology, University of Edinburgh, Edinburgh, EH9 3JT, UK*

Key words: cell-lineage selection, somatic selection, mosaicism, probability of fixation, genetic load

Abstract

Selection within the individual may have played a critical and creative role in evolution, boosting the survival chances of mutations beneficial to the cell and the individual, hindering the spread of deleterious mutations, and reducing the genetic load imposed on the population. We review the literature and present new results to describe the effects of cell-lineage selection on the rate and fixation probability of new mutations. Cell-lineage selection can alter these quantities by several orders of magnitude. Cell-lineage selection is especially important in the case of rare recessive mutations, which are hidden from selection at the individual level but may be exposed to selection at the cellular level. Because selection within the individual acts as a sieve eliminating deleterious mutations and increasing the frequency of beneficial ones, mutations observed among progeny will have been pre-selected and are more likely to increase cell proliferation than would randomly generated mutations. Although many authors have focused on the potential conflict between selection at the cellular and individual levels, it must be much more common that the two levels act concordantly. When selection at the cell and individual levels act in a cooperative manner, increased rather than decreased opportunity for germline selection will be favored by evolution.

Introduction

One of the great puzzles raised by J.B.S. Haldane (1937, 1957) was how could selection, in an organism with thousands of genes, act on even a small fraction of the segregating alleles within a population without leading to an extremely high variance in fitness and an intolerably large number of selective deaths. In a recent publication, G.C. Williams made the pertinent observation that Haldane's dilemma concerning the cost of natural selection has not been solved, merely that investigators have moved on to worry about other problems (Williams, 1992, pp. 143–148). It has generally been assumed that differential fitness is manifest in terms of a decreased survival or fertility of individuals. In this paper, we will explore selection acting at a different level, as differential survival or replication of cells during the development of an individual. We will argue that selection among cells may play a pivotal role in eliminating deleterious mutations and in spreading novel advantageous alleles. The mutation load and the cost of selection may, therefore, be paid in currency not at the individual level but rather at the cellular level.

Selection among cells during the development of an individual has played only a cameo role in population genetics theory. The most compelling argument for ignoring selection within an individual is that all cells within a multicellular individual are recently derived from a common single-celled ancestor (the zygote or spore) and are hence closely related with little expected variation (Maynard Smith & Szathmáry, 1995, p. 244). Mosaicism is, however, commonly observed (Whitham & Slobodchikoff, 1981; Gill & Halverson, 1984; Klekowski, 1988; Gill et al., 1995) and is generated by a number of mechanisms including mitotic mutation, mitotic recombination, and mitotic gene conversion (John & Miklos, 1988). Together these processes generate the genetic variation that makes possible evolutionary change within an individual. Selection within an individual acts upon differences

in cell growth rate and survival due to these genetic changes and can lead to gene frequency change within a generation (Klekowski & Kazainova-Fukshansky, 1984; Slatkin, 1984; Antolin & Strobeck, 1985; Hastings, 1989, 1991; Otto & Orive, 1995).

Genetic changes that occur within strictly somatic cell-lines that have no possibility of producing progeny are evolutionary dead-ends and are not discussed further in this paper (even though many of the results are still applicable). Our focus is instead on cell-lineage or germline selection, by which we mean selection among those cells that may lead to offspring through gametes, spores, fragmentation or any other means. Germline selection occurs when cells that differ genetically (because of mutation, crossing over, or gene conversion that occur during mitosis) also differ in their propensity to proliferate or survive during development. A second type of selection that clearly occurs at the cellular level is gamete selection, caused by competition among gametic cells for fertilization. Although it is likely that gamete selection is common in plants where the haploid genotype is expressed in pollen, the situation is different in higher animals where gametes express only the diploid genotype of their progenitor cells. Thus, in animals, the wide genotypic variation among gametes is not mirrored by a large phenotypic variation and the potential for selection is reduced. Variation created during mitotic expansion of the diploid progenitor cells can, however, be subsequently acted upon by gamete selection even in metazoans (Hastings, 1989).

The focus of the current paper is germline selection, which we shall explore using two quite different models of development that we shall call exponential and turnover. In the exponential model, we assume that there is a growing population of cells created by binary cell division, leading to exponential growth of the developing individual (Otto & Orive, 1995). In the turnover model it is assumed that a distinct protogermline grows to a constant size and then cells merely turn over in a birth/death process prior to their differentiation into gametes (Hastings, 1989, 1991); mutation and mitotic crossovers are assumed to be negligible during the initial expansion of cell numbers. The exponential model is more general and appropriate for 'lower' organisms, which may have short life cycles and/or no clearly defined germline. The turnover model is more appropriate for long-lived metazoans with distinct, differentiated germlines. Although these models assume vastly different developmental programs, we find that the results of the models are qualitatively similar.

Germline selection has two primary effects depending on the source of genetic variation within an individual. When a genetic mosaic is created by mutation, germline selection will alter the frequency of the mutation transmitted to offspring. Mutations increasing cell proliferation will be more likely to be transmitted, while those decreasing cell proliferation will be less likely. The quantity of interest is, therefore, the probability that an offspring carries a new mutation that arose during development. This quantity is necessary to determine the observed mutation rate, the probability of fixation of a novel mutation, and the mutation load within a population.

After a mutation has appeared and is segregating within a diploid population, gene conversion and mitotic recombination within heterozygotes will create genetic mosaics, individuals containing both heterozygous and homozygous cells. Germline selection in these mosaics will alter the transmission properties of the different alleles, in a manner analogous to segregation distortion or meiotic drive. For instance, if germline selection favors the homozygous mutant cell over the heterozygous one, then the mutant will be observed in more than half of the gametes, i.e., in more gametes than expected on the basis of Mendelian segregation. We shall quantify this effect in terms of the proportion, k, of the allele A transmitted by an Aa heterozygote, by analogy to models of meiotic drive (e.g., Crow & Kimura, 1970; Hartl & Clark, 1989). A new allele within a population will change in frequency over time due to its relative fitness advantage among individuals and also due to whatever segregation distortion is caused by germline selection. If we represent the frequency of a rare allele as P ($P << 1$), then, assuming random mating, its frequency in the next generation, P', will be

$$P' = 2k\frac{W}{\overline{W}}P, \tag{1}$$

where W is the fitness of the individual heterozygote scaled by $\overline{W}$, the mean fitness of individuals in the population. With Mendelian segregation, k = 1/2, so that the allele increases whenever the heterozygous fitness is greater than the mean fitness. With k not equal to 1/2, the allele will spread when rare if $P' > P$ or

$$k\frac{W}{\overline{W}} > 0.5 \tag{2}$$

(Prout, 1953; Sandler & Novitski, 1957). Equation (2) demonstrates that an allele can spread (or be eliminated) merely by altering k away from its normal Mendelian value of 0.5, even if it is neutral at the individual level. In many cases, however, variants that are favourable in the germline, such as more efficient 'housekeeping' genes, will also be beneficial for the adult. In this case, we need to determine whether the segregation distortion caused by germline selection will have a substantial effect on the spread of the variant allele. It is also possible that 'driven' alleles (with $k > 1/2$) spread despite the fact that they reduce adult fitness. Therefore, a conflict between selection in the adult and germline may arise as in many meiotic drive systems (Buss, 1982; Michod, 1996). This potential conflict has been cited as one reason why animal gametes express the diploid genotype of the adult rather than their own haploid genotype, thereby minimising the genetic differences between the adult and gametes and the extent of genetic conflict (Maynard Smith & Szathmáry, 1995, p. 175).

In the following pages, we briefly review some of the relevant data and discuss previous models of selection within the individual. We then develop some novel results that focus on the influence of cell-lineage selection on the spread and fixation probability of a new allele. We next revisit the question of genetic loads in the presence of cell-lineage selection. Finally, we conclude by discussing how the benefits of germline selection may have shaped the evolution of sexually dimorphic germlines in large metazoan animals.

Background

For selection to act among cells, cells must differ genetically. While this might seem to impose a rather stringent limitation on cell-lineage selection, the sheer number of cells that must be produced during the development of all but the smallest of organisms ensures that almost every individual is a genetic mosaic. Mosaicism is, unfortunately, well known in humans in the form of cancerous changes that may occur at any time in the DNA content of our cells. That mosaicism is the rule rather than the exception in humans is made clear by the example of retinoblastoma, a disease causing tumorous growth of the retina. Although this disease is recessive at the cellular level, it is dominant at the individual level. This difference is due to the fact that somatic genetic changes almost always occur prior to or during the development of the retina and, in heterozygotes, cause at least one cell to become homozygous for the mutant allele, thereby inducing tumor formation (Cavanee et al., 1983). Genetic mosaicism also plays an extremely positive role in the immune system, underlying the changes in B cells essential to acquired immunity to disease (French, Laskov & Scharff, 1989).

Mosaicism is more difficult to see within the germline of animals, although it undoubtedly occurs. We can, however, infer its existence using a number of lines of evidence. Firstly, *de novo* mutations that cause disease in humans are often paternally derived (Vogel & Rathenberg, 1975; Francke et al., 1981; Winter et al., 1983). Secondly, a comparison of sequence divergence of genes on the sex chromosomes and autosomes (which spend 33% and 50% of their time, respectively, in males) suggest that the latter evolve faster (Miyata et al., 1987). Similarly, Shimmin, Chang, and Li (1993) compared zinc-finger protein genes on the X (ZFX) and Y (ZFY) chromosomes, finding that ZFY (which spends all of its time in males) evolves faster. All of the above phenomena are thought to be due to the larger number of cell divisions that occur in spermatogenesis than in oogenesis. The number of cell divisions would, of course, only have an impact on the mutation rate if mutations occur mitotically (Vogel & Rathenberg, 1975), and so we may take the fact that male and female mutation rates differ as evidence that mutations do occur within the germline creating genetic mosaicism. Interestingly, Miyata et al. (1987) reported a smaller divergence rate in 'general' or 'housekeeping' loci that are expressed in all tissues including the germline (where they would be subject to germline selection), than in those loci which were tissue specific (and hence not generally subject to germline selection). Much more direct evidence for mosaicism within the germline comes from studies of spontaneous mutations in *Drosophila* (Woodruff & Thompson, 1992) and humans (Vogel & Rathenberg, 1975). These studies find that individuals carrying a new mutation tend to have many more mutant siblings than expected by chance, suggesting that the parental germline was mosaic for the mutation.

Mosaicism in plants is well known for its important role in horticulture, having led to such novelties as seedless grapes, navel oranges, and pink grapefruits (Gill, 1987). Somatic mutations have led not just to novelties but to a large proportion of the varieties of plants in cultivation. For example, 5000 of the 8800 plants varieties grown in Europe in 1899 originated as somatic mutations, that were then artificially selected and propagated (Whitham & Slobodchikiff, 1981). The

reader is referred to the excellent reviews by Whitham and Slobodchikiff (1981) and Gill et al. (1995) for further evidence regarding genetic mosaicism in plants.

The extent of mosaicism within an organism depends ultimately on the rate at which mitotic mutation, recombination, and gene conversion occur. Estimates of the mutation rate per gene per individual generation (μ) generally fall between 10^{-7} and 10^{-4} (Voelker, Schaffer & Muhau, 1980; Li & Graur, 1991). The rate of mitotic crossing-over (X) tends to be higher, within the range of 10^{-4} to 10^{-2} per individual generation in *Drosophila* (Gethmann, 1988), 10^{-5} to 10^{-4} per individual generation in plants (Evans & Paddock, 1979), and 10^{-7} to 10^{-5} per cell generation in yeast (Lichten & Haber, 1989; Yuan & Keil, 1990). Although any two cells randomly chosen from an individual are likely to be identical at a locus (Antolin & Strobeck, 1985), the fact that there are so many cells within an individual and so many genes that can mutate implies that substantial genetic variation must exist within an individual. Furthermore, even when a mutation is extremely rare, perhaps only happening once in the history of a population (e.g., a particular insertion, deletion, or rearrangement), when the mutation occurs, it will exist within a mosaic individual and its fate will depend on the form of selection acting upon cells.

The impact that cell-lineage selection can have on the genetic variation within a mosaic individual depends, in turn, on the number of cell divisions per individual generation. The greater the number of cell divisions, the more selection will be compounded within an individual generation. Data from multicellular organisms indicate that the number of cell divisions per generation can be quite large. From zygote to zygote, the number of cell divisions has been estimated to be: 50 for maize (Otto & Walbot, 1990), 35 for *Drosophila* (at 18 days for males and at 25 days for females; Drost & Lee, 1997), 25 for female mice (Drost & Lee, 1997), 62 for 9-month-old male mice (Drost & Lee, 1997), 23 for human females (Vogel & Rathenberg, 1975), and 36 for human males that reproduce at puberty plus 23 per year thereafter (Vogel & Rathenberg, 1975). The number of cell divisions per generation is strongly age dependent in *Drosophila* and in mammalian males, rising, for example, in human males from 36 for a 13-year-old to over 500 for a 35-year-old. In our figures, we will use a range of values for the number of cell divisions per individual generation that span these observations.

Models of cell-lineage selection have been developed under a variety of assumptions concerning development and selection within an organism. Here we briefly review these models (more details may be found in the review by Gill et al., 1995, or in Otto & Orive, 1995).

Luria and Delbrück (1943) were the first to analyze a model of mutation during clonal development. They considered an exponentially growing population of bacteria in which mutations continually arise conferring resistance to a virus. Although the virus was not present during clonal expansion, Luria and Delbrück wished to know what fraction of the population would be resistant to the virus if the virus were later added to the bacterial population. They found that this fraction is extremely variable, reflecting the fact that those cultures in which mutations arose early would have a high fraction of resistant cells (a 'jackpot') while the majority of cultures would have few mutations that appeared fairly late in clonal expansion. In essence, their model can be used to examine gametic selection in an expanding population of germline cells, because selection is absent during development (of the bacterial culture) but acts among cells at the end of development. Similar observations and explanations have been made for *Drosophila* (Woodruff & Thompson, 1992; Woodruff, Huai & Thompson, 1996).

Slatkin (1984) also studied a model that falls under the category of gametic selection. He looked at the probability of fixation of a somatic mutation in trees with multiple branches whose fertility depends on whether or not the mutation is present. Mutant cells did not proliferate at different rates but could have differential reproductive success, depending on the extent to which they altered characteristics of the branch in which they occurred. Slatkin found that such fertility selection among the branches of a tree would increase the probability of fixation of beneficial mutations, especially if the tree is more fertile as a whole due to the presence of the somatic mutation (termed hard selection). Nevertheless, the increase in the probability of fixation was only proportional to the fertility advantage of the mutation, since selection occurred only during reproduction.

More recently, Hastings (1989) analyzed a model of gametic selection in animals, focusing on the change in allele frequency that occurs due to competition among gametes for internal fertilization. He found that gamete selection was 'an efficient means of eliminating unfavorable alleles from the population', but that the effect was weak compared with selection at the individual

level, again because selection acted only once during the individual generation. The fact that gametic selection occurs only at one stage (after development) limits its impact. In contrast, cell-lineage or germline selection occurs throughout development and, therefore, has an effect that accumulates throughout an individual generation. Several authors have explored this phenomenon, using models of development tailored specifically to plants, animals, and organelles.

Modelling development in seed plants, Klekowski and Kazarinova-Fukshansky (1984) investigated the observed number of mutations in organisms that develop from a series of apical meristems. They focused on the fate of mutant cells present in the first apical meristem. They found that cell-lineage selection has the greatest effect on mutation rates when apical meristems are composed of a large number of cells and when initial cells are chosen infrequently during development to form the next apical meristem. In other words, cell-lineage selection is most effective when the proto-germline undergoes very few bottlenecks during development. Using a different model of plant development, Antolin and Strobeck (1985) studied differential growth of the buds on a plant and found that positive selection caused substantial increases in the frequency of mutant buds, especially among long-lived organisms.

Hastings (1991) investigated germline selection within metazoans, assuming a constant population size of replicating cells, a model that we explore further within this paper. This model assumes that the germline grows rapidly to a set size after which time cell death and replication cause turnover, but not growth, of the germline. Hastings concluded that germline selection caused substantial changes in the rate of mutation and spread of mutant alleles: 'under plausible assumptions of germline molecular biology, the mutation rate per gamete may differ up to 100-fold between loci due to selection within the germline' (1991, p. 1171).

Birky (1991) modeled the processes of mutation, selection, and drift among the organelles of a cell. He found that mutations that reduce the replication rate of organelles are unlikely to fix within the cell and tend soon to be lost. Conversely, beneficial mutations are more likely to avoid loss by drift and to fix within the cell. Again these results indicate that selection within an organism (here among organelles) is likely to alter the observed mutation rate of deleterious and beneficial mutations.

Finally, Otto and Orive (1995) developed a variety of models similar to those above to examine the potential influence of somatic selection on the genome-wide mutation rate and deleterious mutation load. The mutation load was found to be very sensitive to changes in the replication rate of cells, especially when many cell divisions occur during development and when development is fairly plastic.

Taken together, these studies demonstrate that within-individual selection can have a major impact on the mutation rate that would be observed among the progeny of an individual. While gametic selection has an effect on mutation rates that is only proportional to the selection coefficient, germline selection of equal strength is compounded throughout development and has a much larger effect on mutation rates. In long-lived organisms, the observed mutation rate may be enhanced by orders of magnitude for mutations that are beneficial to cell proliferation and decreased by orders of magnitude for mutations that are deleterious.

Effects of germline selection on observed mutation rates

The distribution of mutations observed among the offspring of an individual will be shaped, in part, by the process of mitotic mutation and cell-lineage selection. We use the term 'underlying mutation rate' to denote the mutation rate per gamete per locus that would be observed in the absence of selection within an individual and 'observed mutation rate' to denote the rate in the presence of within-individual selection. Germline selection increases the observed mutation rate for alleles that cause cells to survive and replicate at a higher rate and decreases it for alleles that impede cell replication. As discussed in the previous section, this phenomenon has been well studied and we simply illustrate the type of results obtained in Figure 1, which gives the observed mutation rate divided by the underlying mutation rate (Π) for an exponentially growing population of germline cells (see Appendix A for the derivation).

One feature that is evident in Figure 1 is that the observed mutation rate is very sensitive to small changes in the replication rate of mutant cells and the number of cell generations per individual generation. This occurs because the effects of cell-lineage selection are compounded over many cell generations, such that small differences in replication rate can cause a large cumulative impact on the composition of the cell population at the end of development.

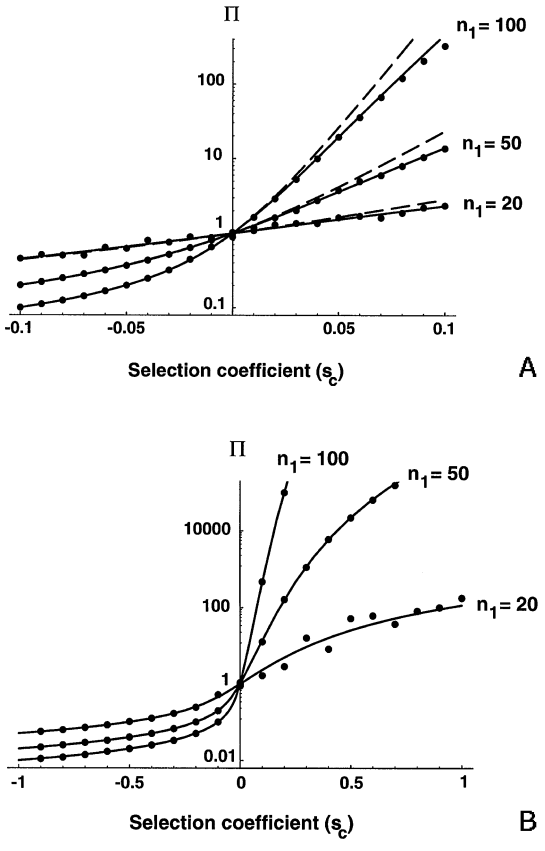

Figure 1. Change in the proportion of adult cells carrying a new mutation due to cell-lineage selection (Π). Simulations (dots) were run of a population of cells growing via binary cell division for n_1 cell divisions, corresponding to an individual generation. Mutations were allowed to accumulate randomly at a small rate (the observed value of Π was independent of the exact rate), causing cells to differ in fitness (abscissa). (a) Weak selection. (b) Strong selection. For comparison, Π is given using equation (A2) (solid curves) and using the approximation (A5) (dashed curves, shown only for weak selection). Note that the observed mutation rate equals Π times the mutation rate in the absence of cell-lineage selection and that the probability of fixation of a new mutation is approximately Π times what it would be in the absence of cell-lineage selection.

Probability of fixation of a new mutant

Cell-lineage selection will have an impact not only on the observed mutation rate but also on the probability that a new mutant will become established within a population. In this section, we estimate the probability of fixation of a mutation that is subject to selection at both the individual and cellular levels. We are especially interested in the fate of a rare mutation, such as a rearrangement, deletion, or insertion in the DNA sequence, that does not occur at a continuous rate within the population. For this class of mutations, the probability of fixation of mutant alleles is a major determinant of the rate of evolution of the population. This phenomenon can easily be studied using the results of Kimura (1962, 1964), who determined the probability of fixation of an allele in a finite population based on a diffusion analysis. In a diploid population of effective size N_e, Kimura found that the probability of fixation of an allele is

$$u = \frac{1 - e^{-4N_e s_I p_0}}{1 - e^{-4N_e s_I}}, \qquad (3)$$

where $1 + s_I$ is the number of offspring carrying the allele per parent with the allele, and p_0 is the initial allele frequency. Kimura's derivation assumes a Poisson distribution of offspring per parent and weak selection, but can be used for both beneficial and deleterious mutations. For a haploid population, N_e must be replaced with $N_e/2$ in equation (3).

Cell-lineage selection has two effects on the probability of fixation of a mutant allele, which we shall explore in turn. The first effect occurs within the genetically mosaic individual created by the mutation itself. Selection within this mosaic individual will alter the initial frequency, p_0, of the mutant allele within the population of reproducing individuals. The second effect occurs in subsequent generations, as genetic mosaics are created by mitotic recombination and gene conversion within heterozygotes. Selection within these mosaic individuals alters the transmission parameter, *k*, and will, therefore, have an effect on the selective spread of an allele (through changes in s_I).

(i) Effects of germline selection on p_0

In Appendix A, we estimate the proportion, *P*, of adult germline cells that are expected to carry a new mutation that occurs sometime during the development of the germline. This proportion is extremely low for a single mutation (equation A2; see Table 1 for notation), since mutations are most likely to arise fairly late in development. Therefore, even if the mosaic individual successfully reproduces, it is unlikely that the mutation will be present in the successful gamete. Consequently, the mutation has an extremely high chance of being lost during the first generation.

We can calculate the probability of fixation by noting that p_0 equals the proportion of gametes within the entire population that will come from the mutant individual and which are mutant, which equals $P/(2N)$ in

Table 1. Parameters of the model with exponential growth in the number of cells. The organism is assumed to grow for a specified amount of time, τ, with cells dividing at a rate dependent on their genotype. Rates of cell division are related to selection coefficients at the cellular level as follows. The fitness of a cell carrying a single mutant is defined as the expected number of daughter cells produced during the amount of time it takes for wild-type cells to produce exactly two daughter cells divided by two (the number of wild-type cells produced in this time). Because, by definition, c_1 is the rate of cell replication per unit time for wild-type cells, a single cell generation takes $1/c_1$ time units, during which time, cells carrying a single mutant produce $2^{c_2/c_1}$ daughter cells. Therefore, the relative fitness of single mutant cells is given by $1 + s_c = 2^{c_2/c_1}/2$. In diploids cells that carry two copies of the mutant allele, the fitness relative to wild-type cells is $1 + \sigma_c = 2^{c_3/c_1}/2$. Note that in standard notation $s = h\sigma$, where *h* is an index of dominance; we use this notation because it is then straightforward to apply our results to haploid organisms (in parentheses) as well as diploids. The mode of gene action in diploids (recessive, additive, or dominant) can easily be recovered by setting h = s/σ

Genotype	Rate of cell division	Number of cell divisions/generation	Fitness at the cellular level	Fitness at the individual level
aa or *(a)*	c_1	$n_1 = c_1\tau$	1	1
Aa or *(A)*	c_2	$n_2 = c_2\tau$	$1 + s_c = 2^{c_2/c_1}/2$	$1 + s_I$
AA	c_3	$n_3 = c_3\tau$	$1 + \sigma_c = 2^{c_3/c_1}/2$	$1 + \sigma_I$

a diploid population. Assuming that *N* is large, the fixation probability of a beneficial allele is approximately $2s_I PN_e/N$. Therefore, the effect of cell-lineage selection on the fixation probability of an allele is proportional to its effect on *P* and is again measured by Π (see Figure 1). Using the estimate for Π obtained in Appendix A for weak selection, cell-lineage selection changes the fixation probability of a mutation by approximately:

$$\Pi \approx \frac{(1 + s_c)^{n_1} - 1}{s_c n_1} \qquad (4)$$

where s_c is the selection coefficient of the allele at the cellular level and n_1 is the number of cell generations per individual generation for non-mutant cells. Equation (4) indicates that the average fixation probability of an allele favorable at the cellular level rises nearly exponentially with the number of cell divisions per individual generation and with the strength of germline selection. Conversely, when cell-lineage selection acts against the mutant allele, its probability of fixation decreases exponentially with n_1 and s_c. Note that the results in Appendix A can also be used to examine the impact of selection acting in opposite directions at the cellular and individual levels.

As an example, consider a slightly beneficial mutation that has a cellular selection coefficient of 0.01. From equation (4), this mutation will have a 30% higher probability of fixation in the presence of germline selection when there are 50 cell divisions per individual generation and 72% higher when there are 100 cell divisions per generation. For strongly selected mutations, the more accurate estimate of *P* given by (A2) must be evaluated numerically. For mutant cells that are twice as fit as non-mutant cells (s_c = 1), the chance of fixation is one million times higher in the presence of cell-lineage selection with 50 cell divisions per generation and 10^{13} times higher with 100 cell divisions per generation.

For deleterious mutations, equation (4) provides a fairly accurate estimate of the effect of cell-lineage selection on the fixation probability for all selection coefficients. With a cellular selection coefficient of -0.01, the probability of fixation of a deleterious mutation is reduced by 21% with 50 cell divisions per generation and by 37% with 100 cell divisions. For a strongly deleterious mutation ($s_c = -1/2$), the chance of fixation is reduced by 96% with 50 cell divisions per generation and by 98% with 100 cell divisions.

In short, by changing the frequency of a new mutation among the gametes of an individual, cell-lineage selection has a substantial influence on the fate of a mutation.

(ii) Effects of germline selection on transmission in heterozygotes

The genetic mosaicism created by the original mutation event is, however, immediately lost in subsequent generations if individuals reproduce via single-celled offspring. In diploid populations, genetic variation among cells is regenerated in subsequent generations by mitotic recombination and gene conversion in heterozygous individuals (Hastings, 1989, 1991; Gill et al., 1995). Because the rate of mitotic crossing-over (*X*) is within the range of 10^{-5} to 10^{-2} per individual gener-

ation, genetic mosaicism generated in heterozygotes may have an important influence on the spread of an allele.

We measure this influence by the proportion, *k*, of *A* gametes transmitted from *Aa* heterozygotes. Changes in *k* from the Mendelian expectation of 1/2 alter the expected number of mutant offspring per mutant parent and affect the chance that a mutation will fix within a population. In an outbred diploid population, a new mutation is nearly always found in heterozygotes and so its overall fitness is $1 + s_I = (1 + s)\,2k$, where s_I measures the expected number of heterozygous offspring per heterozygous parent and *s* is the viability or fertility advantage of a heterozygote individual. If the allele is initially present in a single heterozygous individual and if selection is weak, then the probability that it will ultimately fix is approximately $2\,s_I$ (Haldane, 1927). If the allele is present at a higher frequency or if it is deleterious, then s_I can be used in equation (3) once *k* is determined.

We estimate *k* in Appendix B using the model of an exponentially growing population of dividing cells (equation A10) and in Appendix C using the model of a constant-sized population of cells with cell turn-over (equation A15). These models give quite similar results as long as selection coefficients are small (less than 0.1), and we, therefore, illustrate *k* as a function of cell-lineage selection for the first model only (Figure 2).

It is quite apparent in Figure 2 that the transmission parameter *k* rises only slowly with the strength of cell-lineage selection, unless there are many cell generations per individual generation. Thus, we generally expect the strength of individual level selection (1 + *s*) to outweigh the fitness effects of segregation distortion (2*k*) caused by cell-lineage selection. This point is illustrated in Figure 3, which shows the fixation probability of an additive mutation assuming that selection is equally strong at the cellular and individual levels. Compared to the case in which cell-lineage selection is absent (solid curve), cell-lineage selection substantially increases the fixation probability of a single new mutant allele by increasing its frequency (p_0) in the gamete pool in the first generation (long dashed curves). Changes in *k* as well as p_0 have an almost indiscernible effect on the probability of fixation (short dashed curves) except when the number of cell generations per individual generation is large (100 or more).

Nevertheless, the transmission of a mutation may play an important role in the spread of alleles that have beneficial effects at the cellular level but are neutral or

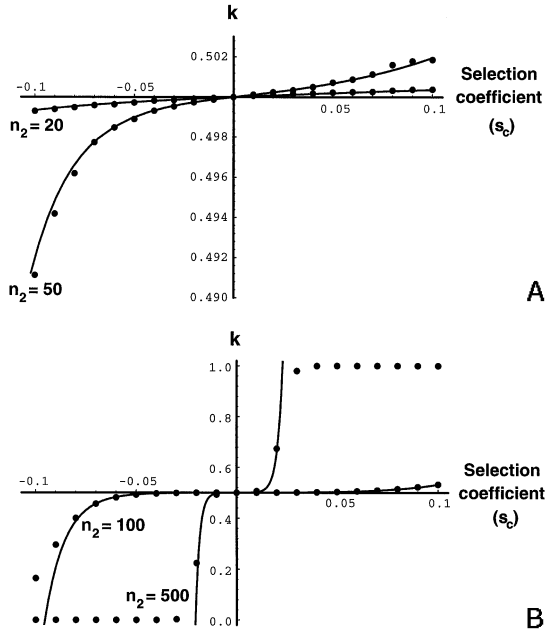

Figure 2. Segregation distortion in heterozygotes due to cell-lineage selection. The proportion of *A* gametes produced by heterozygous individuals is shown as a function of the strength of selection at the cellular level. Simulations (dots) are compared to the expected value of *k* from equation (A10) (solid curves). The rate of conversion per individual generation, *X*, was set to 0.001. The strength of cell-lineage selection favoring *AA* homozygotes was set to twice that favoring *Aa* heterozygotes ($\sigma_c = 2s_c$), i.e., gene action is additive. (a) Few cell generations per individual generation: segregation distortion is weak. (b) Many cell generations per individual generation: segregation distortion is strong.

nearly neutral at the individual level. In this case, s_I in equation (3) becomes 2*k* - 1 and even small increases in *k* above the Mendelian value of 0.5 will increase the probability of fixation of an otherwise neutral allele. A potentially important case in which cell-lineage selection may determine the fate of an allele is with recessive mutations. Such mutations are neutral in the germline in which they arise (i.e., p_0 is unaffected) and in terms of natural selection on heterozygous individuals. They are not neutral in the germlines of heterozygotes, however, since *Aa* cells give rise to *AA* mutant cells through mitotic recombination and selection can act upon the resulting genetic variance. Such selection will increase *k* and will, therefore, make it much more likely that a beneficial recessive allele will fix within the population.

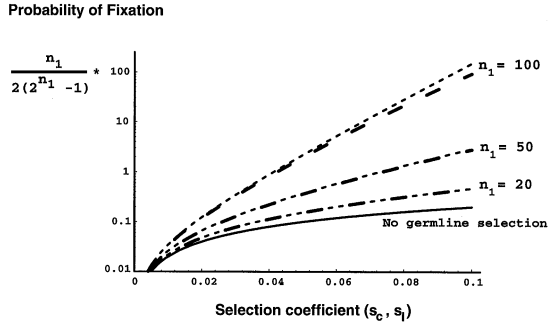

Figure 3. The probability of fixation of a single mutation that occurs once during development. In the absence of cell-lineage selection ($\sigma_1 = 2s_I; \sigma_c = 2s_c = 0$), the initial fraction of mutant gametes in the population is $p_0 = P_{\text{neutral}}/(2N)$ with P_{neutral} given by equation (A3); the solid curve shows the fixation probability obtained from equation (3) using this value for p_0. With cell-lineage selection acting only in the generation in which the mutation appears ($\sigma_c = \sigma_I = 2s_c = 2s_I$), p_0 becomes $P/(2N)$ with P given by equation (A2); the probability of fixation is then shown by long dashed curves. If cell-lineage selection acts in subsequent generations to change k as well as p_0, the fixation probability is hardly affected unless n_1 is large (short dashed curves). The population size, N, was assumed to be large relative to selection ($Ns_c >> 1$). Note that the ordinate is scaled to P_{neutral}, the proportion of cells within the adult individual that are expected to carry a new mutation in the absence of cell-lineage selection.

(iii) Simulation results

We verified the results obtained in Appendices A-C by simulating the two models explored within this paper, that of an exponentially growing population of cells and of a constant-sized population of cells. The first series of simulations assumed that development occurs by a series of binary cell divisions, with the rate of cell division determined by a cell's genotype. Mutations and/or mitotic recombination occurred randomly during each cell division. A run was terminated once the individual had 'developed', i.e., after an appropriate number of cell divisions. This process was repeated 10,000 times (100,000 times for Figure 1B), and the results are shown as dots in Figures 1, 2. The initial zygotic cell was wild-type in Figure 1 and heterozygous in Figure 2. The simulations coincide extremely well with the theoretical predictions.

In the second series of simulations, we tracked a population of cells that underwent birth and death such that the total population size remained constant. For this model, the transmission parameter k can be estimated analytically by assuming an infinite population of cells as in Appendix C (equation A15). In an individual, however, the number of cells will always be finite. To assess the importance of this difference, we used simulations with a total number of cells, ν, varying from 100 to 10,000. Each of these cells underwent, on average, n cell divisions per individual generation. We ignored events that occurred during the initial expansion of the proto-germline leading to the population of ν cells, but followed mutations and conversions that occurred during the rest of the lifespan. During this second phase, a single cell was chosen at random to die and was replaced by the daughter of a randomly chosen cell (a birth-death model). This was repeated ν times every cell generation, for a total of νn times every individual generation. The chance that a cell was selected to replicate was determined by its frequency times its relative fitness, divided by the mean fitness of all cells. Every offspring cell so produced was subjected to mutation and gene conversion at the appropriate rates.

The results shown in Figure 4 demonstrate that equation A15 provides an accurate estimate for k only when the size of the population of cells is large ($\nu \geq 10{,}000$). For smaller numbers of cells, k may be significantly closer to 1/2 than predicted. Although expressed in terms of advantageous alleles, the results also hold for the elimination of deleterious mutations whose transmission is 1 - k. As can be seen in Figure 4a-d, individuals tend to produce either 100% *A* alleles or the normal Mendelian ratio, although this is due in part to the parameters chosen (i.e., fairly strong selection with fitnesses of 1, 1.1, and 1.2 for *aa*, *Aa*, and *AA* cells, respectively). Because selection is quite strong, the spread of *AA* cells throughout the germline population of cells is rapid once an *AA* cell has been created. Increasing the number of cells within the germline (Figure 4) or decreasing the strength of selection results in the *AA* genotype taking longer to spread through the germline, and more individuals transmit *A* alleles with values between 50% and 100% (Figure 4). The observation of intermediate values of k is probably more realistic and is more consistent with the observation of mosaics noted previously (Woodruff & Thompson, 1992, and refs. therein). If, however, selection is strong as in the simulations, an approximation for the average transmission ratio, k', may be found for a finite population of cells (Appendix D). This enables us to illustrate the effects of increasing cell lineage number on k' and on the fixation probability assuming no selection at the individual level (Figure 4e). Although this only applies to strongly selected alleles, it has the advantage of clearly showing the benefits and

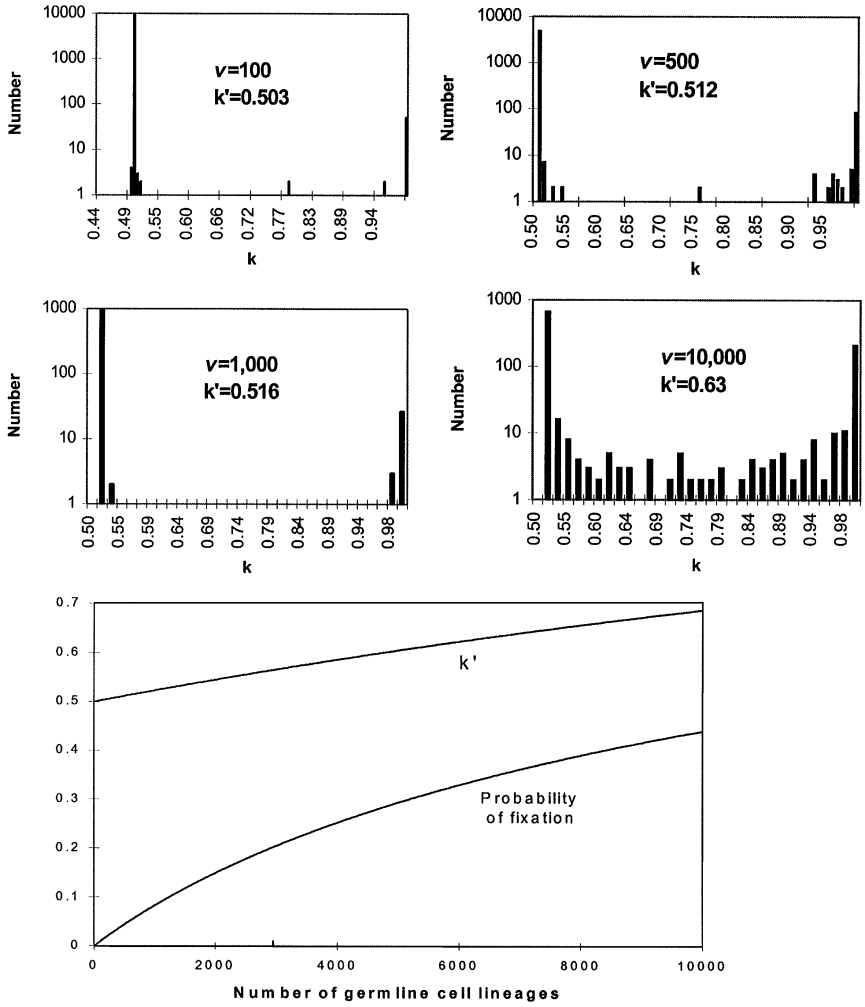

Figure 4. Simulation of germline selection when the number of cells within the germline population is finite i.e., ν = 100, 500, 1,000 or 10,000; other parameters were n = 100, X = 10^{-3}, s_c = 0.1, and σ_c = 0.2. Ten thousand individuals were simulated when ν = 100 or 500 and one thousand when ν = 1,000 or 10,000. The distribution of k for each simulation was obtained and plotted as histograms with the mean value, k', given in the caption; note log scale on the Y axis. For comparison, the deterministic matrix method described in Appendix C (equation A15), which assumes an effectively infinite number of cells within the germline, gives k' = 0.624; so, providing there are thousands of cells turning over in the population, the matrix method provides a good estimate of the segregation ratio. An approximation was derived for the average transmission ratio, k, as a function of the number of cells within the germline population, ν (see Appendix D). Figure 4e shows this relationship between k' and ν as well as the relationship between the probability of fixation of a mutant allele and ν assuming that selection does not act at the individual level (as for recessive mutations).

implications of increasing germline size, which will be discussed later.

Genetic load in a population under mutation/selection balance

We have shown here and in previous work that germline selection can drastically reduce the gametic mutation rate to deleterious alleles (Hastings, 1991; Otto & Orive, 1995). Specifically, the mutation rate at a locus is changed by germline selection from μ to $\Pi\mu$ (equation A6). Because the genetic load is two times the mutation rate in a diploid population without epistasis and with mutations that are not fully recessive (Crow, 1970), the genetic load in a diploid population will be changed by germline selection by an amount Π (see Figure 1). Similarly, since the genetic load is equal

to the mutation rate in a haploid population (Crow, 1970), germline selection will reduce the genetic load of a haploid population by Π, where we would expect Π to be much smaller in haploids since mutations are then fully expressed at the cellular level rather than masked (see Otto & Orive, 1995). Notice that the load will be decreased by germline selection if alleles that are deleterious at the individual level are also deleterious at the cellular level, but will be increased for mutations that increase cell proliferation.

In addition to changing the observed mutation rate, gene conversion and mitotic recombination allow germline selection to act against deleterious mutations in heterozygotes, reducing the frequency of transmission of these mutations. In keeping with Table 1, let 1 - s_I equal the viability of *Aa* individuals and 1 - σ_I the viability of *AA* mutant individuals. The frequency of the deleterious allele *A* in a diploid population at mutation-selection balance will then be

$$\hat{q} = \frac{\Pi\mu}{1-(1-s_I)2k}, \tag{5}$$

where *k* is the transmission of the mutant allele from heterozygotes. [In the absence of germline selection (Π = 1 and *k* = 1/2), equation (5) reduces to the standard result for semi-dominant mutations: $\hat{q} = \mu/s_I$, but a more accurate analysis is needed to estimate $\hat{q}$ for recessive mutations.] The mean fitness of the population will be approximately $1 - 2\hat{q}s_I$ in the presence of semi-dominant mutations and $1 - \hat{q}^2\sigma_I$, with fully recessive mutations. Consequently, the mutation load as defined by Crow (1970) will be

$$\text{Load} = 2\hat{q}s_I = \frac{2\Pi\mu s_I}{1-(1-s_I)2k} \quad \text{for semi - dominant mutations} \tag{6}$$

$$\text{Load} = \hat{q}\sigma_I = \left(\frac{\Pi\mu}{1-2k}\right)^2 \sigma_I \quad \text{for recessive mutations.} \tag{7}$$

Cell-lineage selection acting against deleterious mutations will reduce Π below one and *k* below 1/2, both of which will result in a reduction in mutation load. The mutation load is most dramatically reduced for recessive mutations (from order μ to order μ^2), since cell-lineage selection allows mutations to be eliminated by biased transmission from heterozygotes (k < 1/2) rather than by individual death of homozygous mutant individuals. [The interested reader can refer to Bengttson (1970) who investigated the same problem albeit in a different context].

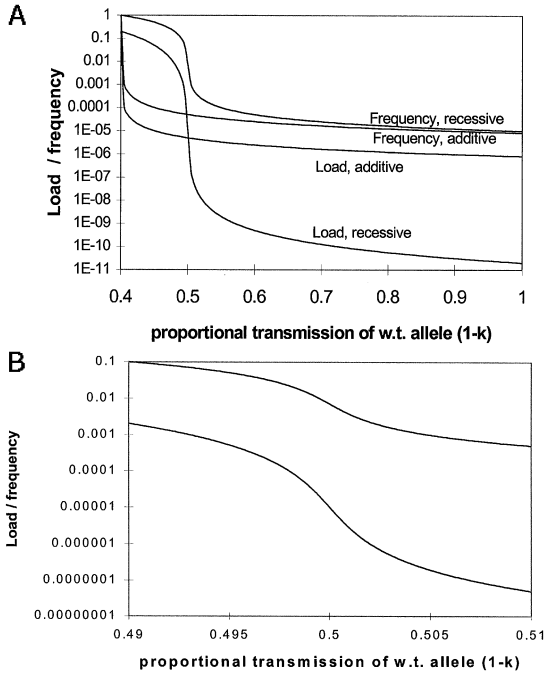

Figure 5. Equilibrium frequencies of deleterious alleles and subsequent genetic load under mutation/selection balance with varying values of *k*. Observed mutation rate was $\Pi\mu = 10^{-5}$, the force of natural selection against the homozygous mutant, σ_I, was 0.2, and the mutations were either additive (s_I = 0.1) or recessive (s_I = 0). (a) the effects of varying 1 - *k* over the range 0.4 to 1.0. (b) The effects of small deviations of 1- *k* around the Mendelian value of 0.5 assuming that the mutants are recessive. Note that *k* measures the transmission of the deleterious mutant, but for consistency with previous figures we plot 1 - *k*, the transmission rate for the wild-type allele.

Figure 5 shows the load that results with sample parameters $\Pi\mu = 10^{-5}$, s_I = 0.1, and σ_I = 0.2. As expected, for both semi-dominant and recessive mutations the equilibrium allele frequency and mutation load drop as the transmission of the wild-type allele (1 - *k*) increases. This effect is particularly striking for recessive mutations, which are eliminated by germline selection within heterozygotes but are invisible to selection at the individual level. As the transmission bias against deleterious mutations increases, the proportion of purifying selection incurred by the germline increases while that in the individual decreases. The results in Figure 5a show that even very small deviations of *k* away from the Mendelian value of *k* = 0.5 can have very large effects on the genetic load imposed by recessive mutations, with the mutation load changing by a factor of 10,000 between transmission ratios of 0.49 and 0.51.

Sexually dimorphic germlines

We have argued that germline selection can be beneficial to the organism, decreasing the observed mutation rate for deleterious mutations while promoting the spread of beneficial mutations. These effects increase with the number of cells within the germline and with the number of cell divisions (Figures 2 and 4). If the advantages to germline selection are so great, why then would small germlines with few cell divisions ever evolve? Human females, for example, produce few gamete cells, with only about 23 cell divisions occurring during oogenesis (Vogel & Rathenberg, 1975). In this section, we discuss the evolution of reduced germlines typically observed in female metazoans and the implications for germline selection.

There are a number of reasons why a diminutive germline might evolve. One is that selection might favor a lowered mutation rate, which could be achieved by reducing the number of mitotic cell divisions that occur within the germline. Another possibility is that selection at the cellular levels and the individual levels are often in conflict, which could favor the evolution of a reduced germline (Michod, 1996, 1997). These two explanations are general in that they predict that both male and female germlines should be small. This is clearly false for mammals. In human males, for example, hundreds of cell divisions may be involved in spermatogenesis (Vogel & Rathenberg, 1975). A straightforward explanation is that intrasexual competition among males for fertilization of females selects for males capable of producing massive numbers of sperm, which is only possible if the male germline undergoes numerous cell divisions. According to this view, the optimal number of germline divisions may be the number observed in females, with sexual selection preventing males from also achieving this optimum. This suggests that sex with males carries an additional cost to females, that of increasing the number of deleterious mutations carried by her offspring (Redfield, 1994).

An alternative explanation that specifically predicts a smaller female germline is that selection favors a lowered mutation rate not in the nuclear genome, but rather in the mitochondrial genome (Short, 1993; Allen, 1996). Mitochondria are particularly susceptible to mutations, because metabolic processes within the organelle create mutagenic, free oxygen radicals and because DNA repair is less efficient in the mitochondrial genome. Furthermore, since mitochondria are inherited maternally and are predominantly asexual in metazoans, selection is less effective at eliminating deleterious mutant alleles once they appear within the mitochondrial population. Therefore, it may be that a metabolically inactive and reduced female germline evolved to prevent mutations from accumulating within the mitochondrial genome. The female nuclear mutation rate may, however, be smaller than the optimal rate for the species. In this case, selection may actually favor an increased male germline as a means of increasing the mutation rate (see e.g., Gillespie, 1981).

Any or all of the above factors may have influenced the evolution of sexually dimorphic germlines. How much does this dimorphism reduce the effectiveness of cell-lineage selection? The answer is 'surprisingly little'. In the dimorphic species that have been studied, most new mutations occur in males. In humans, for example, we would expect roughly ten times more mutations to occur in males than in females based on the number of cell divisions involved in gamete formation, which is consistent with a number of estimates of the male-to-female ratio of mutation rates (Francke et al., 1981; Winter et al., 1983; Shimmin, Chang & Li, 1993). Mutations that arise in males experience many more rounds of cell division and cell-lineage selection than those that arise in females. Therefore, for most new mutations that occur, Π (the observed mutation rate with cell lineage selection divided by the expected mutation rate without it) must be calculated using the number of cell divisions in males. Cell-lineage selection will, therefore, be nearly as effective at decreasing the mutation rate to deleterious alleles and increasing the mutation rate to beneficial alleles as it would if both sexes had large germlines as in males.

There is, however, one effect of cell-lineage selection that is reduced by having a sexually dimorphic germline: germline selection on recessive alleles will be reduced. By creating a transmission bias in their favor, cell-lineage selection promotes the spread of rare, recessive, beneficial alleles, thus substantially increasing their probability of fixation (Figure 4e). Similarly, by reducing the transmission of recessive deleterious mutations, cell-lineage selection reduces the mutation load due to these mutations (Figure 5). With a sexually dimorphic germline, however, there will be a much reduced transmission bias in the sex, generally females, with few germline cell divisions (Figure 2). Therefore, the extent of non-Mendelian transmission (k - 1/2) is reduced by approximately one half when only one sex has a large germline. This effect is explored in Figure 6. We assume that germline selection is ineffective in the female so that $k = 0.5$. Because

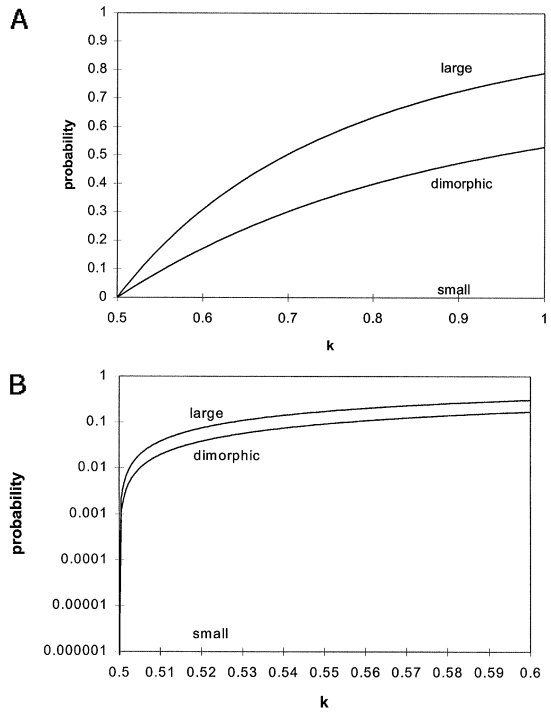

Figure 6. The probability of fixation of a new recessive, beneficial mutation as a function of *k*. Three scenarios are investigated, one where both sexes have a large germline, one where they both have a small germline, and one where the two sexes are dimorphic. In the latter case, normal Mendelian segregation ($k = 0.5$) is assumed to occur in one of them (the 'female' with few germline cell divisions). Since the population size is large, the probability of fixation of a recessive mutation in a population where both sexes have a small germline (resulting in nearly Mendelian transmission) is negligible.

the new mutation is equally likely to appear in a male or a female after its initial appearance, the distribution of the number of offspring is the composite of two equally weighted Poisson distributions, one with $k = 0.5$ (corresponding to transmission from females) and one with *k* variable according to the strength of germline selection (corresponding to transmission from males). The probability of fixation can then be solved using a branching process approach (Haldane, 1927; Crow & Kimura, 1970). Figure 6 shows the effect of sexual dimorphism on the probability of fixation of a new recessive beneficial mutation. As expected, the probability of fixation is nearly halved by the presence of a diminutive female germline.

General discussion

One of the primary benefits of selection within an individual is that it provides a selective sieve, eliminating deleterious mutations and promoting beneficial ones, leading to the 'disproportionate proliferation of those variants favored by environmental demands' (Buss, 1983 p.1390). The models of cell-lineage selection described in this article have demonstrated that selection within an organism can have a major impact on the rate of appearance of new mutations and on their fate. A rare mutation, such as a chromosomal rearrangement, may arise only once in a single cell in a single individual and will have very little chance of ultimately fixing within the population even if it confers a substantial selective advantage upon the individual. Cell-lineage selection can greatly increase the mutation's representation among the reproductive cells of an individual when the beneficial mutation first appears (right hand regions of Figure 1). This can increase the chance that the mutation will become established within the population by orders of magnitude (see also Woodruff, Huai & Thompson, 1996).

Similarly, selection at the level of the cell can eliminate many of the deleterious mutations that arise within a population. This effect can be seen in the left hand regions of Figure 1: the number of deleterious mutations observed in an offspring decreases exponentially with the strength of cell-lineage selection acting on the mutations. The decrease in the observed mutation rate would reduce the mutation load of a population (Hastings, 1991; Otto & Orive, 1995). It would also reduce the susceptibility of a small population to drift load and mutational meltdown (Lynch & Gabriel, 1990). Drift load is caused by the fixation of deleterious alleles in small populations, but this will occur at a lower rate when germline selection as well as individual selection acts against deleterious mutations.

The genetic mosaicism that naturally arises when a mutation first appears is immediately lost, however, when an individual produces offspring through a single cell. Cell-lineage selection would, therefore, be rendered ineffective were it not for a number of mitotic events that create genetic mosaics. Although such events include chromosomal deletions, duplications and rearrangements, mitotic recombination and gene conversion are of particular interest since these processes create a mixture of homozygous and heterozygous cells within individuals that were initially heterozygous and occur frequently (Gethmann, 1988; Evans & Paddock, 1979; Lichten & Haber, 1989; Yuan

& Keil, 1990). Cell-lineage selection in a mosaic of heterozygous and homozygous cells will effectively alter the frequency of mutant alleles among the reproductive cells of an individual, which will result in a transmission bias (i.e., meiotic drive, or, more appropriately, 'mitotic drive'). We investigated the extent of meiotic drive predicted as a result of mitotic recombination and germline selection (see Figure 2). For the parameters investigated, the extent of meiotic drive is fairly weak and will play only a minor role in the spread of an allele unless the number of cell divisions within the germline is large (as in human males) or the strength of selection at the cellular level is much larger than the strength of selection at the individual level. This latter condition necessarily holds in the case of recessive mutations. By definition, these mutations (whether beneficial or deleterious) will experience little selection at the individual level while they are rare. Mitotic recombination will, however, frequently create homozygous recessive cells within a heterozygous individual even when the allele is rare within the population. Any meiotic drive that results from cell-lineage selection would, therefore, govern the selective spread or elimination of the recessive mutation while it is rare. As illustrated in Figures 4 and 5, meiotic drive due to germline selection increases the probability of fixation of beneficial recessive alleles, while decreasing the equilibrium frequency of deleterious recessive alleles and the mutation load that they cause.

Germline selection has several interesting implications for the evolution of mutation rates within a species. Although modifiers of mutation are generally assumed to alter the mutation rate in the same way regardless of the effect of the mutation (Gillespie, 1981; Liberman & Feldman, 1986), a modifier of mutation that acts by altering the strength of cell-lineage selection will increase the observed mutation rate for mutations that are beneficial to cell function and will decrease the observed mutation rate for those mutations that decrease the ability of cells to divide and replicate. Increasing the effectiveness of cell-lineage selection is, therefore, a means by which evolution could reduce the deleterious mutation rate without a concomitant reduction in the production of beneficial mutations.

We have assumed that selection generally acts concordantly at the levels of the cell and the individual. There are, however, a number of important examples in which selection at the cellular level drives the spread of genetic changes that are deleterious to the individual. Cancer is an obvious example. Another is the 'stalkless' mutant in cellular slime molds (Buss, 1983). Stalkless cells gain an advantage by contributing disproportionately to the fruiting structure of the slime mold, but stalkless fruiting bodies are spread less efficiently and are disfavored at the individual level. The existence of such examples of conflict between the cell and the individual is thought to be a potent threat to the integrity of multicellular individuals (Michod, 1996, 1997). The need to reduce such conflict has been invoked to explain the evolution of a distinct germline (Buss, 1987), the expression of the parental diploid genotype in sperm (Maynard Smith & Szathmáry, 1995, p. 175), as well as mechanisms of mutual policing within the individual (Frank, 1995; Michod, 1996).

While examples of conflict are known, it must be far more common that there is cooperation in terms of selection at the cellular and individual levels. This becomes obvious when one considers that the vast majority of mutations that affect fitness at the individual level decrease fitness; certainly they do not all increase fitness at the cellular level, and in many cases they must also limit the ability of cells to grow and divide. We expect that loss-of-function mutations in household genes would generally be detrimental to both the cell and the individual. Similarly, mutations that improve the efficiency of metabolic pathways may often be beneficial at both levels. Even more complex relationships are possible between selection at the cellular and individual levels. For instance, there may be heterozygous disadvantage with respect to individual fitness but directional selection at the cellular level. In this case, biased transmission can drive a population through an underdominant fitness valley provided that $(1 - s)\,2k > 1$.

Much work remains to be done to determine the role of germline selection within evolution. We do not know how much conflict between the levels of selection may occur before the benefits of cooperation outlined in this article are overwhelmed. We do not know how much the observed mutation rate reflects the action of germline selection. We know little about selection acting at the individual level, let alone its relationship to selection at the cellular level. Many fruitful questions remain to be addressed at both the theoretical and experimental levels. The most critical piece of information that we lack is how selection within an individual affects the spectrum of mutations observed among the progeny of that individual. An extremely valuable experiment might compare the distribution of offspring fitnesses from young and old males, where both groups

were irradiated at the same, early stage in development (e.g., as larvae in *Drosophila*). The older males will have undergone more germline cell divisions, during which time cell-lineage selection will have operated (see Gaul, 1958, for a similar experiment in barley). Such experiments are needed to determine whether cell-lineage selection should continue to play only a cameo role in evolutionary biology or should be given equal billing with individual selection.

Acknowledgements

We thank Maria Orive and Michael Whitlock for many insightful discussions about somatic selection and Joni Drost for providing critical information regarding the number of cell divisions per individual generation. We are also grateful to Troy Day, Michael Whitlock, and an anonymous reviewer for their helpful comments on the manuscript. SPO was sponsored by the Natural Sciences and Engineering Research Council of Canada. IMH was supported by the Medical Research Council of Britain.

Appendix A. Expected proportion of adult cells that carry a new mutation

In this model, we assume that development occurs by binary cell division, creating an exponentially growing population of cells. We consider a new mutation that occurs in tissue leading to reproductive cells, whether or not this tissue is confined to a germline. We assume that the rate of cell division is dependent on the genotype of the cell, creating the potential for selection within an individual. Cells carrying the new mutation have a relative fitness of $1 + s_c$, which is assumed to be the heterozygous fitness if the organism is diploid. The mutant allele may either increase ($s_c > 0$) or decrease ($s_c < 0$) the rate of cell division. Parameters of the model are further described in Table 1.

First, consider a single mutation that occurs sometime during development. Assuming that the initial cell was non-mutant, the probability that the mutation occurs in one of the 2^x cells present at cell generation *x* is

$$p_x = \frac{2^x}{\sum_{i=1}^{n_1} 2^i} = \frac{2^x}{2(2^{n_1} - 1)}. \quad \text{(A 1)}$$

If the mutation happens in cell generation *x* (after x/c_1 time units have passed), the resulting number of mutant cells in the adult will be $m_x = 2^{c_2(\tau - x/c_1)}$. Similarly, the total number of wild-type cells will be $w_x = (2^x - 1)2^{c_1(\tau - x/c_1)}$, since $(2^x - 1)$ cells will not have mutated and will divide at rate c_1 from then on. The expected proportion of mutant cells in the adult (*P*) can thus be calculated as

$$P = \sum_{x=1}^{n_1} p_x \frac{m_x}{w_x + m_x}. \quad \text{(A 2)}$$

The actual distribution of mutant cells in models such as this one is extremely skewed (Luria & Delbrück, 1948; Lea & Coulson, 1949; Figure 4 of this paper). Mutations are unlikely to occur early in development, but when they do the proportion of mutants is very high. More often, mutations happen late in development leading to a very small proportion of mutant cells. We are interested, however, in the average probability of fixation, and so the expected value of *P* given by (A2) will serve our purposes.

When $c_1 = c_2$, equation (A2) reduces to

$$P_{\text{neutral}} = \frac{n_1}{2(2^{n_1 - 1})}, \quad \text{(A 3)}$$

which conforms to our expectation that extremely few cells should be mutant, since the mutation is likely to occur late in development. With cell-lineage selection ($c_1 \neq c_2$), equation (A2) can be evaluated numerically (Mathematica code available upon request). An approximate solution may be obtained, however, if it is assumed that the total number of cells in the adult is relatively unchanged by the processes of cell mutation and selection. The proportion of mutant cells in the adult is then:

$$P \approx \frac{n_1}{2(2^{n_1} - 1)} \frac{(1 + s_c)^{n_1} - 1}{s_c n_1}. \quad \text{(A 4)}$$

This approximation was tested within the range of plausible values for n_1 (1-500 cell divisions per individual generation) and found to be fairly accurate whenever $-1 \leq s_c \leq 0.05$.

Cell-lineage selection, therefore, changes the proportion of adult cells carrying a mutation by approximately:

$$\Pi = \frac{P}{P_{\text{neutral}}} \approx \frac{(1 + s_c)^{n_1} - 1}{s_c n_1}. \quad \text{(A 5)}$$

If mutations recur at a rate μ per individual generation (or approximately μ/n_1 per cell division), then the proportion of mutant cells in the adult becomes *P* times the expected number of mutations that arise throughout development, or

$$P_\mu = P \frac{\mu}{n_1} \sum_{i=1}^{n_1} 2^i = P_{\text{neutral}} \Pi \frac{\mu}{n_1} 2(2^{n_1} - 1) = \mu\Pi, \quad \text{(A 6)}$$

which is Π times higher in the presence of cell-lineage selection. Equation (A6) assumes that each new mutation is independent in its effect, which is reasonable as long as $\mu\Pi$ is much smaller than one. This condition is easily satisfied for mutations at a single locus; for genome-wide mutations, the methods of Otto and Orive (1995) should be used.

These estimates were tested against simulations, and the results are shown in Figure 1. Π obtained numerically using equation (A2) provides an excellent fit to the simulation results (solid lines), while the approximation (A5) is satisfactory only when $-1 \leq s_c \leq 0.1$ (dashed lines). Figure 1 shows clearly that the proportion of mutant cells in an adult is strongly affected by the process of cell-lineage selection during development.

Appendix B. Segregation in heterozygotes

Mitotic recombination or gene conversion in a heterozygote produces daughter cells that are homozygous at a locus. We assume that there is no bias in conversion so that *aa* and *AA* homozygotes occur with equal frequency. We let *X* equal the rate of conversion per locus per individual generation, so that X/n_2 is the rate of conversion per cell generation given that there are n_2 cell divisions in a heterozygote. We wish to estimate the proportion, *k*, of gametes carrying

allele *A* produced by an *Aa* individual. Our derivation is analogous to that used in Appendix A. We first consider a single conversion event that occurs sometime during development and that converts a heterozygous cell into a homozygous one.

The probability, p_x, that the conversion will occur in one of the 2^x cells present at cell generation *x* is given by

$$p_x = \frac{2^x}{\sum_{i=1}^{n_2} 2^i} = \frac{2^x}{2(2^{n_2}-1)}. \quad \text{(A 7)}$$

With probability 1/2, this conversion produces a wild-type *aa* cell and otherwise it produces a mutant *AA* cell. If the conversion happens in cell generation *x* (after x/c_2 time units have passed) and creates an *aa* cell, the expected number of *aa* in the adult would be $w_x = 2^{c_1(\tau - x/c_2}$. If it creates an *AA* cell, the expected number of *AA* cells in the adult would be $m_x = 2^{c_3(\tau - x/c_2)}$. In either case, the expected number of heterozygous cells in the adult is $h_x = (2^x - 1)2^{c_2(\tau - x/c_2)}$. From these estimates, we can calculate the expected proportion of *A* alleles in the adult due to a single conversion event:

$$k_{\text{single}} = \sum_{x=1}^{n_1} p_x \left(\frac{1}{2}\left(\frac{1/2h_x}{w_x + h_x} \right) + \frac{1}{2}\left(\frac{m_x + 1/2h_x}{m_x + h_x} \right) \right), \quad \text{(A 8)}$$

which we write as $k_{\text{single}} = 1/2 + \kappa$. As long as homozygous cells remain rare, the total number of conversion events expected per individual will be

$$\frac{X}{n_2} \sum_{i=1}^{n_2} 2^i = \frac{X}{n_2} 2(2^{n_2} - 1), \quad \text{(A 9)}$$

and the expected proportion of *A* alleles in the adult will be

$$k = \frac{1}{2} + \kappa \frac{X}{n_2} 2(2^{n_2} - 1). \quad \text{(A 10)}$$

Equation (A10), which may be evaluated numerically, estimates the amount of segregation distortion that we expect to see in heterozygous individuals due to gene conversion and cell-lineage selection.

In Figure 2, we compare the amount of segregation distortion observed in simulations to that estimated by equation (A10), finding an excellent fit. This figure demonstrates that the amount of segregation distortion that is expected is small unless selection at the cellular level is strong (s_c greater than 0.1 in magnitude) or there are a large number of cell divisions per generation (100 or more). In addition, the amount of segregation distortion depends directly on *X* and will be smaller than shown if the rate of mitotic crossing over and gene conversion is below 0.001 per individual generation.

Appendix C. A deterministic model of germline selection

In this Appendix, we analyze the turnover model describing a germline cell population that remains relatively constant in size, but which turns over continually due to cell death and replication. Further details about the method may be found in Hastings (1989, 1991). The relative frequencies of the three germline genotypes are held in a vector **F**, whose elements $\mathbf{F}_1$, $\mathbf{F}_2$ and $\mathbf{F}_3$ correspond to *aa*, *Aa* and *AA*, respectively. The probability that a cell of genotype *j* is created from a cell of genotype *i* during a single mitotic cell division is given by the transition matrix $\mathbf{T}_{i,j}$,

$$T = \begin{bmatrix} 1 - 2\frac{\mu}{n} & 2\frac{\mu}{n} & 0 \\ \frac{\mu_r}{n} + \frac{X}{2n} & 1 - \frac{\mu}{n} - \frac{\mu_r}{n} - \frac{X}{n} & \frac{\mu}{n} + \frac{X}{2n} \\ 0 & 2\frac{\mu_r}{n} & 1 - 2\frac{\mu_r}{n} \end{bmatrix}, \quad \text{(A 11)}$$

where μ is the mutation rate from *a* to *A*, μ_r is the reverse rate from *A* to *a*, *X* is the rate of mitotic crossing over (all per individual generation), and *n* is the number of cell generations per individual generation. Mutation is assumed to be sufficiently rare that double mutations in a single cell generation may be ignored. Selection acts every cell generation, with the fitness of a cell given by the diagonal elements of a fitness matrix, **W**:

$$W = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 + s_c & 0 \\ 0 & 0 & 1 + \sigma_c \end{bmatrix} \quad \text{(A 12)}$$

If there are *n* mitotic cell divisions per individual generation, then the three genotypes in the adult are held in a vector **F[n]** given by **F[n]** = **F[0](TW)**^{n}. Since we do not divide by the mean fitness each generation, the elements of **F[n]** do not sum to unity so we renormalize this vector to obtain the frequency of each genotype.

When the individual is initially wild-type (*aa*), **F[0]** is given by {1,0,0}. Solving for **F[n]** and renormalizing gives:

$$F[n] \approx \left\{ 1 - 2\frac{\mu}{n}\frac{(1+s_c)^n - 1}{s_c}, 2\frac{\mu}{n}\frac{(1+s_c)^n - 1}{s_c}, 0 \right\} \quad \text{(A 13)}$$

to leading order in μ. The frequency of *A* gametes will equal the frequency of *Aa* cells divided by two, which is identical to that obtained from equation (A6) with (A5). The prediction of this model is, therefore, exactly the same as the prediction based on an expanding population of germline cells when selection is weak.

When the individual is initially heterozygous (*Aa*), **F[0]** is given by {0,1,0}. Assuming that X >> μ, μ_r, **F[n]** is proportional to:

$$\text{F[n]} \approx \left\{ \begin{array}{l} \frac{X}{n}\frac{\left(\left(\frac{1}{1+s_c}\right)^n - \left(1 - \frac{X}{n}\right)^n\right)}{2\left(\frac{1}{1+s_c} + \frac{X}{n} - 1\right)}, \left(1 - \frac{X}{n}\right)^n, \\ \frac{X}{n}\frac{\left(\left(\frac{1+\sigma_c}{1+s_c}\right)^n - \left(1 - \frac{X}{n}\right)^n\right)}{2\left(\frac{1+\sigma_c}{1+s_c} + \frac{X}{n} - 1\right)} \end{array} \right\}, \quad \text{(A 14)}$$

and the proportion of *A* gametes transmitted by a heterozygote is equal to:

$$k = \frac{F_3[n] + \frac{F_2[n]}{2}}{F_1[n] + F_2[n] + F_3[n]} \quad \text{(A 15)}$$

which is approximately:

$$k \approx \frac{1}{2} + \frac{X}{n}\frac{(\sigma_c - s_c + s_c(1+\sigma_c)^n - \sigma_c(1+s_c)^n)}{2s_c(\sigma_c - s_c)(1+s_c)^{n-1}} \quad \text{(A 16)}$$

to leading order in *X*. Evaluating (A16) and (A10) numerically has shown that the two estimates of *k* are nearly identical when selection is weak. The differences that appear when selection is strong are due to the fact that mutations are guaranteed to occur early in this deterministic model, but are unlikely in the model of a growing population of cells.

Appendix D. Algebraic approximation of k as a function of the number of cells

In this Appendix, we develop an approximation for *k* for the case of a finite population of cells, which are originally heterozygous. We assume that an *AA* cell created by a crossover event always becomes fixed in the germline if it is not eliminated by chance (100% transmission of *A*). Otherwise, we assume that 50% of transmitted alleles are *A*. These assumptions are realistic for the parameters used in Figure 4.

An *AA* cell that appears during development will have a probability of ultimate fixation within the population of ν cells equal to:

$$u_c = \frac{1 - e^{-(\sigma_c - s_c)}}{1 - e^{-v(\sigma_c - s_c)}} \qquad \text{(A 17)}$$

(see Ewens, 1979, equation 3.51). Note that Ewens' definition of *s* is twice the selective difference between genotypes *Aa* and *AA*, or 2 (σ_c - s_c) in our notation, and that the number of cells in the germline (ν) replaces his 2N.

The total number of conversion events that will occur within the population of cells is νX per individual generation, but only half of these create *AA* cells. As long as *X* is small, we can approximate the number of *AA* cells created by conversion during development by a Poisson distribution with mean $\nu X/2$. The probability that either no *AA* cells are created during development or that all *AA* cells created fail to fix within the germline is, therefore:

$$f(0) \approx \sum_{j=0}^{\infty} \frac{e^{-vX/2}(vX/2)^j}{j!}(1-u_c)^j = e^{-vXu_c/2} \qquad \text{(A 18)}$$

using u_c from (A17). Thus, an approximate value for the proportion of *A* gametes transmitted by heterozygotes is

$$k = \frac{f(0)}{2} + (1 - f(0)) = 1 - \frac{f(0)}{2}. \qquad \text{(A 19)}$$

This formulation assumes that selection is strong enough or the life cycle long enough that those cells that would ultimately fix within the population of cells rise in frequency fast enough to approach fixation during the lifespan of the individual.

k estimated from equation (A19) compares favorably to simulation results with σ_c = 0.2, s_c = 0.1, and *X* = 0.001 (see Figure 4):

	Algebraic approximation		Simulation	
ν	f(0)	*k*	f(0)	*k*
100	0.995	0.502	0.99	0.503
500	0.976	0.512	0.98	0.512
1000	0.954	0.523	0.97	0.516
10000	0.621	0.689	0.69	0.63

where f(0) in the simulations was obtained by summing classes between 0.45 and 0.55. For large ν , the deterministic equation A15 provides a more accurate estimate for *k* (0.624).

In the absence of selection at the individual level, the fixation probability of a mutation which is initially present in a single individual can be found by numerically solving the branching process with offspring distributions given by two Poisson distributions with means 0.5 and *k* weighted by $f(0)$ and 1 - $f(0)$ respectively (Figure 4e). Alternatively, the diffusion equation (3) with s_I = 2*k* - 1 can be used as a first order approximation to the fixation probability.

References

Allen, J.F., 1996. Separate sexes and the mitochondrial theory of ageing. J. Theor. Biol. 180: 135–140.

Antolin, M.F. & C. Strobeck, 1985. The population genetics of somatic mutations in plants. Am. Nat. 126: 52–62.

Bengttson, B.O., 1970. The effect of biased gene conversion on the mutation load. Genet. Res. 55:183–187.

Birky, C. W., Jr., 1991. Evolution and population genetics of organelle genes: mechanisms and models, pp. 112–134 in Evolution at the Molecular Level, edited by R.K. Selander, A.G. Clark and T.S. Whittam. Sinauer Associates, Inc., Sunderland, MA.

Buss, L.W., 1982. Somatic cell parasitism and the evolution of somatic tissue compatibility. Proc. Natl. Acad. Sci. USA 79: 5337–5341.

Buss, L.W., 1983. Evolution, development, and the units of selection. Proc. Natl. Acad. Sci. USA 80: 1387–1391.

Buss, L.W., 1987. The Evolution of Individuality. Princeton University Press, Princeton, NJ.

Cavanee, W.K., T.P. Dryja, R.A. Phillips, W.F. Benedict, R. Godbout, B.L. Gallie, A.L. Murphree, L.C. Strong & R.L. White, 1983. Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. Nature 305: 779–784.

Crow, J.F., 1970. Genetic loads and the cost of natural selection, pp. 159–168 in Mathematical Topics in Population Genetics, edited by K. Kojima. Springer-Verlag, New York

Crow, J F. & M. Kimura, 1970. An Introduction to Population Genetics Theory. Harper & Row, New York

Drost, J. & W. Lee, 1998. The developmental basis for germline mosaicism in mice and *Drosophila.* Genetica: 102/103: 421-443.

Evans, D.A. & E.F. Paddock, 1979. Mitotic crossing-over in plants, pp. 315–351 in Plant Cell and Tissue Culture, Principles and Applications, edited by W.R. Sharp, P.O. Larsen, E.F. Paddock, and V. Raghavan. Ohio State University Press, Columbus, OH.

Ewens, W.J., 1979. Mathematical Population Genetics. Springer-Verlag, New York.

Francke, U., R.M. Winter, D. Lin, B. Backay, J.E. Seegmiller & W.L. Nyhan, 1981. Use of carrier detection tests to estimate male to female ratio of mutation rates in Lesch-Nyhan disease, pp. 117–130 in Population and Biological Aspects of Human Mutation, edited by E.B. Hook & I.H. Porter. Academic Press, London.

Frank, S.A., 1995. Mutual policing and repression of competition in the evolution of cooperative groups. Nature 377: 520–522.

French, D L., R. Laskov & M.D. Scharff, 1989. The role of somatic hypermutation in the generation of antibody diversity. Science 244: 1152–1157.

Gaul, H., 1958. Present aspects of induced mutations in plant breeding. Euphytica 7: 275–289.

Gethman, R.C., 1988. Crossing over in males of higher Diptera (Brachycera). J. Hered. 79: 344–350.

Gill, D.E. & T.G. Halverson, 1984. Fitness variation among branches within trees, pp. 105–116 in Evolutionary Ecology, edited by B. Shorrocks. Blackwell Scientific Publications, Oxford.

Gill, D.E., 1987. Intraplant genetic variability (topic of a roundtable discussion), pp. 269–275 in Series Entomological (The Hague) V. 41. Insects - Plants, edited by V. Labeyrie, G. Fabres, and D. Luchaise. Kluwer Academic Publishers, Boston MA.

Gill, D.E., L. Chao, S.L. Perkins & J.B. Wolf, 1995. Genetic mosaicism in plants and clonal animals. Annu. Rev. Ecol. Syst. 26: 423–444.

Gillespie, J.H., 1981. Mutation modification in a random environment. Evolution 35: 468–476.

Haldane, J.B.S., 1927. A mathematical theory of natural and artificial selection. V. Selection and mutation. Proc. Camb. Phil. Soc. 23: 838–844.

Haldane, J.B.S., 1937. The effect of variation on fitness. Am. Nat. 71: 337–349.

Haldane, J.B.S., 1957. The cost of natural selection. J. of Genetics 55: 511–524.

Hartl, D.L. & A.G. Clark, 1989. Principles of Population Genetics. Sinauer Associates Inc., Sunderland, MA.

Hastings, I.M., 1989. Potential germline competition in animals and its evolutionary implications. Genetics 123: 191–197.

Hastings, I.M., 1991. Germline selection: population genetics of the sexual/asexual lifecycle. Genetics 129: 1167–1176.

Hastings, I.M., 1994. Selfish DNA as a method of pest control. Philosophical Transactions of the Royal Society 344: 313–324.

John, B. & G.L. Miklos, 1988. The Eukaryotic Genome in Development and Evolution. Allen & Unwin, London.

Kimura, M., 1962. On the probability of fixation of mutant genes in a population. Genetics 47: 713–719.

Kimura, M., 1964. Diffusion models in population genetics. J. Appl. Prob. 1:177–232.

Klekowski, E.J., Jr., 1988. Mutation, Developmental Selection, and Plant Evolution. Columbia University Press, New York.

Klekowski, E.J., Jr. & N. Kazarinova-Fukshansky, 1984. Shoot apical meristems and mutation: Selective loss of disadvantageous cell genotypes. Amer. J. Botany 71: 28–34.

Lea, D.E. & C.A. Coulson, 1949. The distribution of the numbers of mutations in bacterial populations. J. Genet. 49: 264–285.

Li, W.-H. & D. Graur, 1991. Fundamentals of Molecular Evolution. Sinauer Associates, Inc., Sunderland, MA.

Liberman, U., & M.W. Feldman, 1986. Modifiers of mutation rate: A general reduction principle. Theor. Pop. Biol. 30: 125–142.

Lichten, M. & J.E. Haber, 1989. Position effects in ectopic and allelic mitotic recombination in *Saccharomyces cerevisiae*. Genetics 123: 261–268.

Luria, S. & M. Delbrück, 1943. Mutations of bacteria from virus sensitivity to virus resistance. Genetics 28: 491–511.

Lynch, M. & W. Gabriel, 1990. Mutation load and the survival of small populations. Evolution 44: 1725–1737.

Maynard Smith, J. & E. Szathmáry, 1995. The Major Transitions in Evolution. W. H. Freeman and Co., Oxford.

Michod, R.E., 1996. Cooperation and conflict in the evolution of individuality. II. Conflict mediation. Proc. R. Soc. Lond. B. 263: 813–822.

Michod, R.E., 1997. Cooperation and conflict in the evolution of individuality. I. Multilevel selection of the organism. Am. Nat. 149: 607–645.

Miyata, T., H. Hayashida, K. Kuma, K. Mitsuyasa & T. Yasunaga, 1987. Male-driven molecular evolution: A model and nucleotide sequence analysis. Cold Spring Harbor Symp. Quant. Biol. 52: 863–867.

Prout, T., 1953. Some effects of variations in the segregation ratio and of selection on the frequency of alleles under random mating (Appendix to L. C. Dunn). Acta Genet. 4:148–151.

Otto, S.P. & M.E. Orive, 1995. Evolutionary consequences of mutation and selection within an individual. Genetics 141: 1173–1187.

Otto, S.P. & V. Walbot, 1990. DNA methylation in eukaryotes: Kinetics of demethylation and *de novo* methylation during the life cycle. Genetics 124: 429–437.

Redfield, R.J., 1994. Male mutation rates and the cost of sex for females. Nature 369: 145–147.

Sandler, L. & E. Novitski, 1957. Meiotic drive as an evolutionary force. Am. Nat. 91: 105–110.

Shimmin, L.C., B.H.-J. Chang & W.-H. Li, 1993. Male-driven evolution of DNA sequences. Nature 362: 745–747.

Short, R.V., 1993. Why sex?, pp. 3–22 in The Differences between the Sexes, edited by R.V. Short & E. Balaban. Cambridge University Press, Cambridge.

Slatkin, M., 1984. Somatic mutations as an evolutionary force, pp. 19–30 in Essays in Honour of John Maynard Smith, edited by P. J. Greenwood, P. H. Harvey and M. Slatkin. Cambridge University Press, Cambridge.

Voelker, R.A., H.E. Schaffer & T. Muhai, 1980. Spontaneous allozyme mutations in *Drosophila melanogaster*: Rate of occurrence and nature of the mutants. Genetics 94: 961–968.

Vogel, F. & R. Rathenberg, 1975. Spontaneous mutations in man. Adv. Hum. Genet. 5: 223–318.

Whitham, T.G. & C.N. Slobodchikoff, 1981. Evolution by individuals, plant-herbivore interactions, and mosaics of genetic variability: The adaptive significance of somatic mutations in plants. Oecologia 49: 287–292.

Williams, G.C., 1992. Natural Selection: Domains, Levels and Challenges. Oxford University Press, Oxford.

Winter, R.M., E.G.D. Tuddenham, E. Goldman & K.B. Matthews, 1983. A maximum likelihood estimate of the sex ratio of mutation rates in Haemophilia A. Hum. Genet. 64: 156–159.

Woodruff, R.C. & J.N. Thompson Jr., 1992. Have premeiotic clusters of mutation been overlooked in evolutionary theory? J. Evol. Biol. 5: 457–464.

Woodruff, R.C., H. Huai & J.N. Thompson Jr., 1996. Clusters of identical new mutation in the evolutionary landscape. Genetica 98: 149–160.

Yuan, L.-W. & R.L. Keil, 1990. Distance-independence of mitotic intrachromosomal recombination in *Saccharomyces cerevisiae*. Genetics 124: 263–273.