Lectures on Mathematics in the Life Sciences
Volume 25, 1994

The Role of Deleterious and Beneficial Mutations in the Evolution of Ploidy Levels

SARAH P. OTTO

October 8, 1993

ABSTRACT. All sexual organisms experience a haploid and a diploid phase during their life cycle, yet the extent of each phase is remarkably variable. We explore the evolution of a locus that alters the timing of meiosis, hence altering the proportion of the life cycle spent in either the haploid or diploid phase. Evolution of the life cycle will occur in response to viability selection acting at fitness loci even when selection does not directly act on the locus modifying the life cycle. Both selection against deleterious mutations and for beneficial mutations are considered. It is found that even when diploids mask deleterious alleles and reveal beneficial ones, the diploid phase of the life cycle will not necessarily increase. Conditions under which each phase is expected to increase are developed. These theoretical results may, in part, explain observed life cycle variation.

1. Introduction

Nearly a century has passed since the recognition of chromatin as the hereditary material underlying evolution (elucidated in parts by Weismann, Hertwig, Kölloker, Strasburger, Sutton, Boveri, and others [**29**]). In 1883, van Beneden found that the number of chromosomes halved during the production of gametes only to double again during fertilization in *Ascaris bivalens*, a threadworm [**29**], thus showing that the number of chromosomes does not remain constant throughout a life cycle. Cytological studies in the 1890's by Strasburger, Guignard, Overton, Farmer and others [**40**] extended this work to various species of plants. These studies revealed the relationship between chromosome number and the alternation of generations, with a reduction division (meiosis) generally

---

1991 *Mathematics Subject Classification.* Primary 92D15; Secondary 65C20, 15A42.

The author was supported by a Miller Post-Doctoral Fellowship and by NIH grant number GM40282 to Montgomery Slatkin.

The final version of this paper will be submitted for publication elsewhere.

0075-8523/94 $1.00 + $.25 per page

leading to the gametophytic (haploid or x) stage and syngamy leading to the sporophytic (diploid or $2x$) stage. By the early part of the twentieth century, it was understood that some organisms are dominated by the x stage (examples among the green algae), some are dominated by the $2x$ stage (gymnosperms and angiosperms) and yet others are characterized by the extensive development of both $2x$ and x tissue (bryophytes, ferns) [**28, 40**]. This information was immediately placed in an evolutionary context, as exemplified in the following quote:

> In the vegetable kingdom evolution seems to have been accompanied by a gradual increase of the $2x$-generation, and a corresponding reduction of the x-generation in point of importance.
>
> – R. H. Lock, 1906, p. 271.

Perhaps typically, Lock saw the increase in diploidy as he surveyed from the "lowly marine organisms and passing upwards... to the flowering plants," as evidence for an advantage to diploidy. He identified a possible advantage emanating from the genotypic variability produced by the union of two distinct genomes within an individual:

> ...it is only in [$2x$ organisms] that the operation of Mendel's law can lead to the production of new combinations of parental characters in the body which represents the main stage of the life history; and that this circumstance may possibly lead to a greater power of adaptability to external circumstances.
>
> – R. H. Lock, 1906, p. 275.

During this century, many have followed Lock in a belief that diploids tend to be more genetically variable and are consequently favored by evolution [**3, 7, 8, 41**]. One form of this argument is that diploids are able to combine alleles into the heterozygous condition and thus have a broader range of genotypes to explore. Another argument is that diplonts (organisms with syngamy immediately after meiosis) produce genetically diverse gametes whereas haplonts (organisms with meiosis immediately after syngamy) produce genetically identical gametes and this might enhance the competitive advantage of diploid life cycles when variability is favorable [**3**]. Lastly, it has been thought that deleterious recessive mutations, maintained at a higher frequency in diploids, can be an important source of genetic variation in the face of new environmental conditions [**22, 37, 38**].

Others have focused directly on the selective consequences to an individual of its ploidy level, rather than focusing on the extent of variation found in haploids or diploids. For example, selection occasionally favors heterozygotes and this type of selection will, to the extent of its prevalence, favor diploidy [**4, 6, 16**]. Another aspect of diploidy that is thought to be to its favor is that rare deleterious mutations will almost always be accompanied by normal alleles in a diploid individual [**38**]. These normal alleles frequently compensate for a good portion of the deleterious effects of a mutation [**39**], which are then said to be masked. This masking of mutations has generally been thought to favor diploidy [**20, 36**].

Complementing these theories that overdominant selection and selection against deleterious mutations favor diploidy, Paquin and Adams argued that selection for beneficial mutations would favor diploids, since diploids have twice as many genes which can mutate to new favorable alleles [**34**].

There is a tendency among evolutionary biologists to use complexity as an appraisal of the evolutionary status of an organism [**37**]. Thus, much attention has been focused on finding the evolutionary advantage possessed by diploids over haploids. The persistence and ubiquity of haploid forms, which overwhelm diploids in sheer number, belies this focus on the evolution of diploidy. The real task ahead is to identify the factors that favor changes in the ploidy level of an organism (whether by extending the haploid or diploid phase) and those that favor the maintenance of a given ploidy level. Why is it that we see closely related species of algae [**3**] and yeast [**14**] with dramatically different life cycles? Why are there organisms that maintain both a haploid and a diploid phase to their life cycles in the face of variants that would allow dominance of one or the other phase [**9, 10, 21**]? While more experimental work is keenly needed, we are also in need of theoretical guidance. Only recently have models addressed the masking hypothesis [**27, 32, 33, 36**], the overdominance hypothesis [**16**], and hypotheses related to the maintenance of both haploidy and diploidy [**23**]. The results have often been surprising, with common wisdom failing under scrutiny. Masking has been found to favor *haploidy* under conditions of low genetic mixing within a population [**32, 33**]. Selection with one allele favored in the haploid phase and a different allele favored in the diploid phase does not support the maintenance of both phases within a life cycle unless special conditions hold[**23**].

In this work, it will be shown that deleterious and beneficial mutations can impart an advantage to diploidy *or to haploidy.* Specific conditions are outlined under which we expect an increase in one or the other phase of a life cycle. To obtain these conditions, we use a model that closely mimics known life cycles, with individuals passing through and experiencing selection both as haploids and as diploids. The results for deleterious mutations are consistent with those obtained by Otto and Goldstein [**32**] for a different life cycle. We then present the first theoretical analysis of the effects of beneficial mutations on ploidy levels. With beneficial mutations, diploids are more likely to bear new advantageous alleles. Selection is more effective, however, in haploids, so that beneficial mutations tend to sweep faster through the more haploid members of a population, imparting an advantage upon haploidy. Which effect is more important depends on the specific selective and genetic conditions under consideration.

2. Methods

We wish to understand the evolution of ploidy levels. In wholly or partially asexual populations, the evolution of sex will necessarily entail changes in ploidy levels of the population at certain stages (see articles by Michod and Gayley and

by Kondrashov in this collection). In this paper, however, we wish to isolate selection for haploidy or diploidy from selection for sexual or asexual reproduction. We thus assume a sexual life cycle or at least a life cycle with a constant proportion of sexuality to asexuality. We then postulate the existence of a gene (the ploidy locus, C), the effect of which is to hasten or delay the timing of meiosis relative to that of syngamy and zygote formation (Figure 1, [**23**]). The proportion of time spent in each phase of the life cycle depends on the genotype (C_iC_j) of the diploid zygote at this locus, according to Table 1. Genetic changes at this ploidy locus can modify the life cycle by either a small amount (micromutations) or dramatically; for instance, it is possible that a single mutation changes an organism from a haplont to a diplont. For mathematical convenience, we assume non-overlapping generations with all individuals in the population producing gametes at the same time regardless of their genotype. Changes in the frequency of an allele at the ploidy locus are assumed to occur indirectly, in response to the selective regime acting at the second locus (the viability locus, A). The survival rate of an individual per generation (or the fertility rate, assuming individual- and not couple-dependent fertility selection) depends, by assumption, only on its genotype at the viability locus according to Table 2. For any particular chromosome, the survival rate per generation must be adjusted according to the proportion of a generation spent in the diploid phase (d_{ij}) and in the haploid phase ($1 - d_{ij}$). For instance, consider an aC_1 chromosome within an AaC_1C_1 diploid: the probability that it will survive through the diploid phase of selection is $(1 - hs)^{d_{11}}$; the haploid aC_1 offspring then produced will survive with probability $(1 - s)^{1-d_{11}}$. Notice that, if d_{ij} equals one, nascent C_iC_j individuals experience selection only as diploids (or as haploids if $d_{ij} = 0$). The ploidy and viability loci recombine at meiosis with a probability of r ($0 \leq r \leq 1/2$). Mutations occur at a rate, $\mu_1/2$, from the wildtype (A) allele to the mutant allele (a) at the viability locus during meiosis and at a rate of $\mu_2/2$ during gametogenesis (see Figure 1; $\mu_1 = \mu_2 \equiv \mu$ will generally be assumed where the total mutation rate per cycle is μ). The results do depend quantitatively on the time during the life cycle when mutations occur[1] (especially when s is large); the qualitative results that will be discussed, however, are independent of the placement of mutation. The mutation process described allows mutations to occur during the reproduction of both haploids and diploids; this not only appears to be more realistic but also provides results that are intermediate between those when mutation is placed at meiosis or at syngamy.

We first examine the effects of deleterious mutations (mutant fitnesses less than one) on evolution at the ploidy locus and then the effects of beneficial mutations (mutant fitnesses greater than one).

[1]For example, Jenkins and Kirkpatrick, this volume, place mutation at meiosis only and obtain slightly different results.

Genotype	Proportion of time spent as a Diploid	Haploid
C_1C_1	d_{11}	$1-d_{11}$
C_1C_2	d_{12}	$1-d_{12}$
C_2C_2	d_{22}	$1-d_{22}$

TABLE 1. The determination of ploidy level.

Genotype	Viability
AA	1
Aa	$(1 \pm hs)^{d_{ij}}$
aa	$(1 \pm s)^{d_{ij}}$
A	1
a	$(1 \pm s)^{d_{ij}}$

TABLE 2. Viability selection at the A/a locus. Selection is positive (+) for beneficial mutations and negative (−) for deleterious ones.

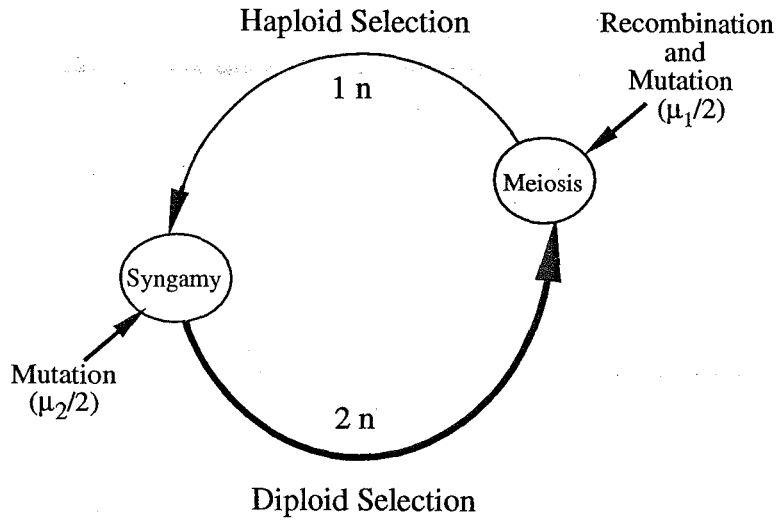

FIGURE 1. Life cycle with alternation of generations. Selection occurs in both the haploid and the diploid phase. The extent of selection in each phase depends on the genotype at a locus (the ploidy locus) which controls the timing of meiosis. Meiosis may occur immediately following syngamy, immediately before it, or any time in between.

3. Results

Following the chromosomal frequencies from one generation to the next, the recursions given in the Appendix A may be derived. These recursions form the basis for the following results.

3.1. Deleterious Mutations. Examining models of individual selection, Perrot, Richerd and Valero [**36**] observed that the masking of deleterious mutations would favor the evolution of diploidy, assuming random mating and free recombination between the loci. Otto and Goldstein [**32**] pointed out that masking has the second effect of allowing mutations to persist and reach higher frequencies in diploids (see also [**15**] and [**33**]). Under conditions of low genetic mixing within a population (low recombination, frequent assortative mating, selfing or asexual reproduction), the chance that a diploid will carry a mutation is so much higher than the probability that a haploid will carry one that individual selection favors the evolution of haploidy, even though diploids mask the mutations that they do carry.

With selection against the a allele ($s > 0$) and with the C_1 allele fixed at the ploidy locus, the recursions in Appendix A approach a polymorphic equilibrium where mutation and selection are balanced:

$$\frac{\mu[1+(1-s)^{(1-d_{11})}]}{2[1-(1-s)^{(1-d_{11})}(1-hs)^{(1-d_{11})}]} \tag{3.1}$$

A new ploidy allele increases in frequency from this equilibrium if the leading eigenvalue determined from the recursions is greater than one (see Appendix B). The leading eigenvalue, equation (B.2), depends on all of the variables of the model in a complicated manner. A few observations may be drawn from a close examination of the leading eigenvalue:

- Whether the leading eigenvalue is greater than or less than unity does not depend on the mutation rate, but the extent of the departure from one is proportional to μ.
- The value of the eigenvalue increases monotonically with r when the new ploidy allele increases the extent of diploid selection.
- The value of the eigenvalue decreases monotonically with r when the new ploidy allele increases the extent of haploid selection.
- When $r = 0$, haploidy is always favored.
- As r is increased from zero, the parameter space in which diploidy is favored increases, with diploidy being favored most often when mutations are highly recessive (h near zero), as illustrated in Figure 2.

Each of these results is completely analogous to results derived in [**32**]. Thus we see that linkage favors the evolution of haploidy while extensive recombination favors the evolution of diploidy. The major departure from previous results is that the invasion criterion is now sensitive to the extent of diploidy among the resident population (d_{11}), as illustrated in Figures 3-5. Above each curve are

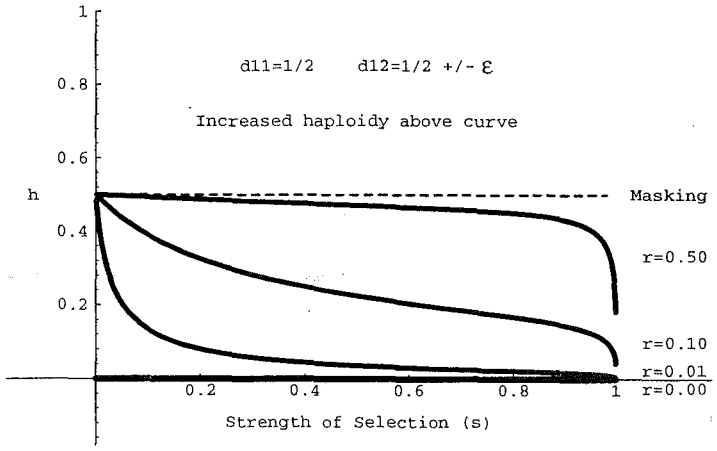

FIGURE 2. Parameters favoring the evolution of increased haploidy (above curve) or increased diploidy (below curve) for various recombination rates. Note that masking occurs in the entire region below the dashed line ($h < 1/2$). Here and elsewhere, ϵ refers to a small quantity.

the values of h and s for which haploidy is favored; below each curve diploidy is favored. The results are clearly sensitive to the exact ploidy alleles under consideration. Consider the point marked by a $*$ ($h = 0.4, s = 0.9$) with free recombination. An initially haploid population ($d_{11} = 0$, Figure 3) cannot be invaded by a ploidy allele that increases the diploid phase by a small amount ($d_{12} = \epsilon$), or by an intermediate amount ($d_{12} = 1/2$), but can be invaded by an allele that leads to a large diploid phase ($d_{12} = 1$). An initial population that experiences both phases equally ($d_{11} = 1/2$, Figure 4) can be invaded by alleles that cause either complete diploidy ($d_{12} = 1$) or complete haploid ($d_{12} = 0$), but can only be invaded by micromutations that increase the extent of diploidy ($d_{12} = 1/2 + \epsilon$). An initially diploid population ($d_{11} = 1$, Figure 5) cannot be invaded by any alleles. Graphical analyses indicate that populations with intermediate values of d_{ij} are more susceptible to invasion and that the initial increase of extreme values of d_{ij} (near zero or one) is supported over a larger parameter range than the invasion of intermediate values. In short, alleles that maintain both phases of the life cycle in intermediate amounts are evolutionarily unstable in this model (in the EGS sense [**11**]); the dominance of either the haploid phase *or* the diploid phase is expected given the appearance of sufficiently diverse modifier alleles. These results thus shed no light on the maintenance of alternating generations, but rather make its occurrence more puzzling.

Despite the quantitative differences between the results obtained using the life cycle illustrated in Figure 1 and that used by Otto and Goldstein [**32**], the qualitative result that genetic associations develop which favor haploidy when

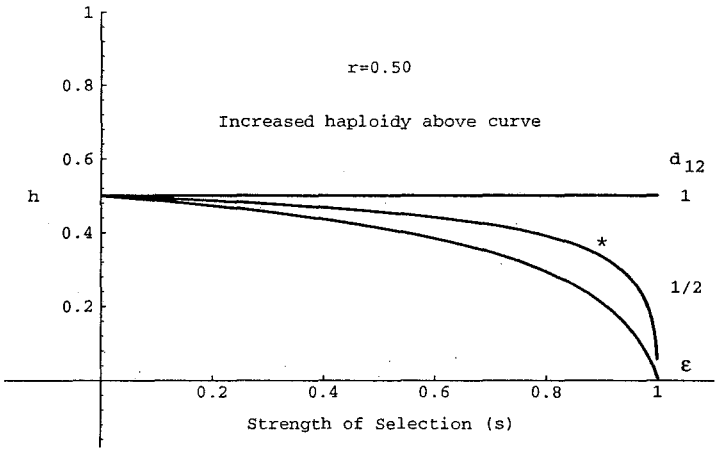

FIGURE 3. Sensitivity to initial ploidy level: starting population is haploid, d_{11}=0. For different values of d_{12}, the curve is given below which the rare allele with the new ploidy level can invade, thus increasing the extent of the diploid phase. Above the curves, the resident ploidy allele is stable to invasion. An $*$ is placed where $s = 0.9$ and $h = 0.375$ for discussion within the text.

recombination rates are low is common to both. These associations develop because individuals bearing mutations are more likely to die if they are haploid since the mutation is then unmasked. Those haploid individuals that do not die but survive selection are less likely to carry mutations and are less likely to bear offspring with mutations. From the opposite perspective, masking permits diploids carrying mutations to survive selection but it thereby allows those mutations to persist among the offspring of these diploids. Consequently, the frequency of deleterious mutations becomes higher among individuals with longer diploid phases than among individuals with longer haploid phases, an association that favors haploidy.

3.1.1. *Comments about non-random mating.* Otto and Marks [**33**] found that selfing, assortative mating, parthenogenesis and other forms of asexual reproduction favor the evolution of haploidy by limiting the genetic mixing that separate haploid (diploid) ploidy alleles from the viability alleles that have recently experienced haploid (diploid) selection. I investigated these forms of non-random mating for the life cycle currently under analysis, finding once again that their effect was to strengthen the extent of genetic associations thereby increasing the parameter range in which haploidy was favored. For any particular organism, mating patterns must thus be understood before we can predict whether haploidy or diploidy would be favored by selection against deleterious mutations.

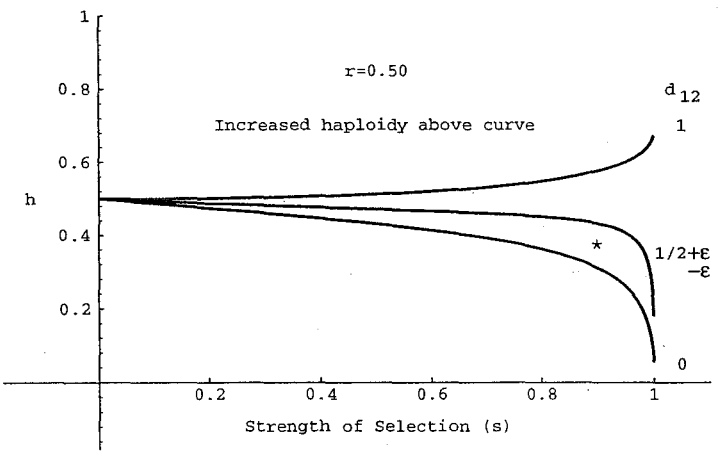

FIGURE 4. Sensitivity to initial ploidy level: starting population is haplont-diplont, d_{11}=1/2. For different values of d_{12}, the curve is given below which the rare ploidy allele can invade if $d_{12} > d_{11}$ but not if $d_{12} < d_{11}$. Above the curves, the new ploidy allele can invade if $d_{12} < d_{11}$ but not if $d_{12} > d_{11}$. An $*$ is placed where $s = 0.9$ and $h = 0.375$ for discussion within the text.

3.2. Beneficial Mutations. Paquin and Adams [**34**] argued that the number of beneficial mutations experienced by a diploid population would be larger than that of a haploid population because twice as many genes would be able to mutate. Their arguments were based on several assumptions: that the populations of haploids and diploids do not interbreed, that the number of haploids would be less than twice the number of diploids, and that there would be no significant differences in the probability that a beneficial mutation would become fixed within a haploid versus a diploid population. We evaluate the influence of sweeps of beneficial mutations on the evolution of ploidy levels in the context described by Paquin and Adams, with non-interbreeding asexual populations, and then in the context of an interbreeding population of sexual haploids and diploids as in Figure 1.

3.2.1. *Asexual haploid and diploid populations.* We wish to know when the rate of adaptation, defined as the rate of accumulation of beneficial mutations, will be faster in a diploid asexual population than in a separate haploid one. While diploids have twice as many mutations within their doubled genome, these new mutations arise in the heterozygous state and are generally less advantageous than if they arose unmasked in a haploid. In comparing the rate of adaptation of separate haploid and diploid populations, then, we must consider both the increased number of mutations in diploids and the decreased advantage of these mutations.

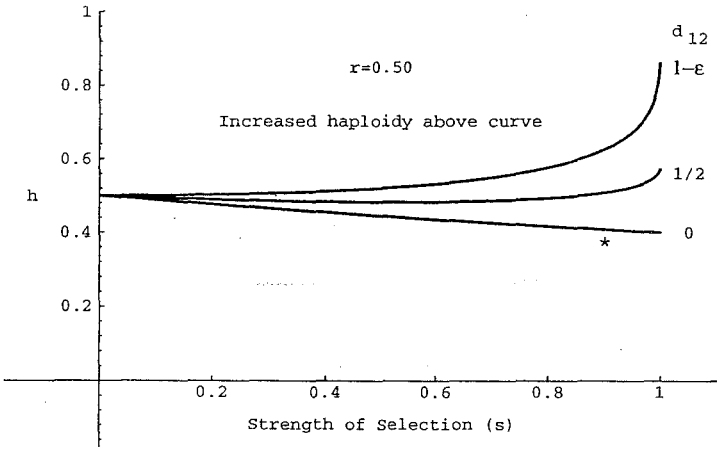

FIGURE 5. Sensitivity to initial ploidy level: starting population is diploid, d_{11}=1. For different values of d_{12}, the curve is given above which the rare allele with the new ploidy level can invade, thus increasing the extent of the haploid phase. Below the curves, the resident ploidy allele is stable to invasion. An $*$ is placed where $s = 0.9$ and $h = 0.375$ for discussion within the text.

In asexual populations, beneficial alleles can only accumulate within a genome if successive mutations occur within a single lineage; mutations occurring in different individuals cannot, in the absence of sex, be recombined into one individual [**30**]. If mutations are infrequent such that a mutation arises and fixes before the next beneficial mutation occurs, then these new mutations will always occur in the lineage with all previous beneficial mutations (technically, all those that survived loss when rare). Here, the nesting of beneficial mutations is ensured. On the other hand, if mutations are frequent, new mutations can occur in different lineages that are present concurrently, but only one of these lineages will ultimately survive. All new mutations that do not occur within the "lucky" lineage will ultimately be lost, even though they impart a selective advantage to their carriers. As a concrete example, consider a new mutation that occurs within a population with a previous beneficial mutation segregating at a frequency of 5%. Most likely, this new mutation will not occur within an individual bearing the previous mutation; all individuals will thus carry no mutations or a single mutation. The single mutant with the highest fitness advantage will eventually fix and the other will be lost from the population. If mutations are very frequent, then, many mutations will simply be lost by competitive exclusion. In this case, the rate of adaptation is generally limited not by the rate of mutation, but by the rate of spread of beneficial alleles (that is, by their selective advantage). The faster good alleles spread, the sooner new mutations can be nested within the

same lineage. In short, the rate of fixation of beneficial mutations in asexual populations will tend to be limited by one of two factors: the mutation rate to new advantageous alleles (when mutations are rare) or the time required for beneficial mutants to reach appreciable frequencies (when mutations are frequent).

In this section, we will determine quantitatively how the rate of adaptation depends on mutation rates and selection coefficients. These results will be discussed in light of expected differences between haploid and diploid asexual populations. We will then reexamine the data of Paquin and Adams with *Saccharomyces cerevisiae*, from which they infer more sweeps of beneficial mutations in diploid populations.

The following analysis is similar in structure to that presented by Kirkpatrick and Jenkins [**26**] and by Weiner et al. [**43**]. Consider an asexual population of size N with beneficial mutations occurring at a genome-wide rate equal to U. ***If*** no other mutations are segregating within the population, a new beneficial mutation that arises will have a probability of eventually becoming fixed that is approximately equal to 2σ, where $1 + \sigma$ is the relative fitness of an individual carrying this new mutation and is assumed to be near one [**19**]. Thus the rate that new beneficial mutations (destined to fix) will arise in an asexual population is:

$$\rho_m = N\ U\ 2\sigma. \tag{3.2}$$

We may use equation (3.2) to compare the rate of adaptation of separate haploid and diploid populations. Let N_H and N_D be the haploid and diploid population sizes, respectively. Assume that beneficial mutations occur at a rate μ per gene per generation, and that there are L genes that may experience beneficial mutations in haploids but twice as many, $2L$, in diploids ($U = \mu L$ for haploids, but $2\mu L$ for diploids). Further, let the relative fitness of an individual with a new beneficial mutation be $1 + s$ in haploids and $1 + hs$ in diploids. Realistically, s and h are random variables following some unknown distribution; for simplicity, we will assume that s and h are constant. We will also assume that h tends to be less than one, presuming that an advantageous allele that occurs within an individual carrying another allele at the same locus (i. e. a diploid) will be less effective. Equation (3.2) thus becomes:

$$\rho_m^{hap} = N_H\ \ \mu L\ \ 2s \qquad \text{for the haploid population and} \tag{3.3}$$

$$\rho_m^{dip} = N_D\ 2\mu L\ \ 2hs \qquad \text{for the diploid population,} \tag{3.4}$$

which assume that each beneficial mutation sweeps through the population before the next mutation appears. For haploid and diploid populations of equivalent size, the rate of adaptation in a diploid population will be $2h$ times that of a haploid population. The diploid population will fix dominant mutations ($h > 1/2$) at a faster rate, but will accumulate beneficial recessive changes ($h < 1/2$) at a slower rate.

When beneficial mutations arise frequently, as they would in very large populations, the probability that any new mutation will survive to fix in the population is reduced from the value of 2σ, since these mutations may arise in individuals that do not have the highest fitness within the population. The faster beneficial alleles sweep through the population, however, the more likely it is that the fittest genotype will be at a high frequency, and the more likely new mutations will be nested within this lineage. To tell whether beneficial mutations tend to occur while previous mutations are still segregating, we now give the time required for a beneficial mutation to fix. In an asexual population of size N, a beneficial mutation that arises with frequency p_1 and selective advantage σ will reach fixation in $\bar{t_f}(p_1)$ generations, on average [**25**], where

$$\bar{t_f}(p_1) \approx \frac{1}{\sigma(1-e^{-2N\sigma})} \int_{p_1}^{1} \frac{(e^{2N\sigma x}-1)(e^{-2N\sigma x}-e^{-2N\sigma})}{x(1-x)} dx \tag{3.5}$$

if p_1 is very small (to translate the model used by Kimura and Ohta into this asexual model, their selection coefficient was doubled and their population size halved). For N sufficiently large, $\bar{t_f}(p_1)$ may be approximated by the deterministic analog calculated as in section 2.1 of [**12**],

$$\begin{aligned} \bar{t_f}(p_1,p_2) &\approx \int_{p_1}^{p_2} \frac{1+\sigma x}{\sigma x(1-x)} dx \\ &= \frac{1}{\sigma}(ln[p_2(p_2-1)^{-(1+\sigma)}] - ln[p_1(p_1-1)^{-(1+\sigma)}]), \\ &= \frac{1}{\sigma}(ln[\frac{p_2}{p_1}] - (1+\sigma)ln[\frac{1-p_2}{1-p_1}]), \end{aligned} \tag{3.6}$$

where p_2 is the gene frequency assigned as the fixation point (recall that an infinite population would take an infinite amount of time to fix completely). If we measure the time that an allele currently in one individual would be in every individual of the population except one ($p_1 = 1/N$ and $p_2 = 1-1/N$), equation (3.6) becomes:

$$\bar{t_f}(1/N, 1-1/N) \approx \frac{(2+\sigma)}{\sigma} ln[N-1]. \tag{3.7}$$

For large N and small σ, we see from equation (3.7) that the time to fixation is roughly proportional to $1/\sigma$ (with the rate of fixation, $\rho_f \equiv 1/\bar{t_f}$, being proportional to σ). In short, if the rate of adaptation is limited not by the appearance of mutations but rather by their spread, we might expect that the rate of accumulation of beneficial mutations would be proportional to the selective advantage of the alleles.

Applying these results to a comparison of haploid and diploid populations, let σ equal s for haploid mutants and hs for diploid mutants, as before. Using equation (3.7), we see that the rates at which beneficial alleles fix in haploid and

diploid populations, respectively, are:

$$\rho_f^{hap} \equiv \frac{1}{\bar{t_f}^{hap}(1/N_H, 1-1/N_H)} \approx \frac{s}{(2+s)ln[N_H-1]}, \tag{3.8}$$

$$\rho_f^{dip} \equiv \frac{1}{\bar{t_f}^{dip}(1/N_D, 1-1/N_D)} \approx \frac{hs}{(2+hs)ln[N_D-1]}. \tag{3.9}$$

In large populations of equal size with relatively weak selection, the rate of fixation of beneficial mutations will be less in diploids by an amount approximately equal to h.

In summary, if mutations are rare and tend to fix before the appearance of the next mutation (the rate of fixation is faster than the rate of mutation: $\rho_f \gg \rho_m$), then from equations (3.3) and (3.4) we would expect a diploid population to accumulate dominant beneficial mutations more rapidly and a haploid population to accumulate recessive beneficial mutations more swiftly. If beneficial mutations occur sufficiently often (the rate of fixation is slower than the rate of mutation: $\rho_f \ll \rho_m$), however, multiple alleles will segregate simultaneously. Here, we expect that haploid populations would adapt faster since beneficial alleles spread faster in haploid populations (with $h < 1$) and thus more nesting of advantageous alleles within a lineage can occur. For further analytical comparisons, see [**31**].

For illustration, let us consider the data of Paquin and Adams [**34, 35**]. They studied populations of *S. cerevisiae*, where population sizes were about 4.9×10^9 and 4.5×10^9 for the haploid and diploid populations, respectively ([**35**], Adams pers. comm.). Paquin and Adams followed changes in marker frequencies to determine when positive selective sweeps were passing through the populations. They found that the rate of sweeps per cell per generation was approximately 3.6×10^{-12} for haploid cells and 5.7×10^{-12} for diploid cells. That is, approximately 60% more sweeps were observed per diploid cell. These figures imply that a beneficial sweep occurred on average once every 57 generations within the haploid populations and once every 39 generations within the diploid populations.

At first glance, this appears to be a significant advantage. As noted by Paquin and Adams and by Charlesworth [**5**], however, this advantage would be partially offset by the fact that haploid populations tend to have higher densities, which could, in nature, translate to larger haploid population sizes. The haploid populations in the chemostat experiments of Paquin and Adams were 40% more dense than the diploid populations. If the haploid population had been 40% larger than the diploid population, the measured advantage to the diploid population would have been reduced, presumably, from 60% to about 14%.

Another source of uncertainty in the interpretation of the experimental results is the quantity L, the number of loci subject to beneficial mutations. It need not be the case that the same number of loci are subject to beneficial mutations in haploid and diploid populations, even when the haploids and the diploids are isogenic since they are not isomorphic. Diploid cells, for instance, tend to

		$\sigma = 0.09$	$\sigma = 0.10$
Stochastic:	$\bar{t_f}(1/N)$	467	422
Deterministic:	$\bar{t_f}(1/N, 1-1/N)$	519	469
Deterministic:	$\bar{t_f}(1/N, 0.05)$	215	194

TABLE 3. Time until fixation in generations. Parameters were chosen from the experiments of Paquin and Adams: $N = 5 \times 10^9$, $\sigma = 0.09$ for diploid mutants, and $\sigma = 0.10$ for haploid mutants. For comparison, in the last row, the times required for a mutation to reach a frequency of 5% are given.

be larger with smaller surface area/volume ratios and tend to be less efficient at nutrient uptake [**1, 42**]. Thus, observing more adaptive sweeps in a diploid population may simply mean that the diploid population was less well adapted to begin with.

Paquin and Adams also estimated the fitnesses of the new beneficial alleles, finding on average that $s = 0.1$ and $h = 0.9$. In Table 3, we apply equations (3.5) and (3.6) to find the expected times to fixation for these mutations ([**12**], p. 61, presents a similar analysis). Clearly, beneficial mutations with the selective coefficients measured take several hundreds of generations to fix and spend a large portion of this time rare. These fixation times are much too large to correspond to the events (the sweeps) measured by Paquin and Adams, which occurred roughly every 50 generations. Therefore, multiple mutations must have been segregating simultaneously, which is consistent with later interpretations that the populations were polymorphic [**2**]. If, however, the populations were subject to such frequent mutations, then we might expect the populations with the faster rate of spread of beneficial alleles (the *haploid* populations) to adapt faster (compare equations (3.8) and (3.9)), in direct contrast to the conclusions in [**34**].

A computer program (available upon request) was developed to refine these arguments. Again for illustration, the parameters of Paquin and Adams were employed. In the simulations, one locus is a marker locus subject to mutations at a rate of 10^{-7}. The remainder of the genome is summarized by the number, n, of beneficial mutations carried relative to the current wildtype, a number which could vary from zero (in the current wildtype) up to eight. Selection favors individuals according to the number of beneficial mutations, with their relative fitness equalling $(1 + \sigma)^n$. Note that only the number of mutations matters and not their location, so that individuals within the same class may have different genotypes but are equally fit. When beneficial mutations are rare, they occur stochastically upon each fitness class according to the frequency of that class. That is, if a number chosen randomly from a uniform distribution between zero and one is less than the product of the mutation rate, the class frequency and the population size, then a mutation occurs moving one individual

(with frequency of $1/N$) from the given class (i) to the next highest fitness class ($i+1$). When beneficial mutations are common (more than one mutation expected in the population for a given fitness class), they occur deterministically. When the current wildtype reaches a low frequency (0.0001), it is grouped with the class carrying one beneficial mutation and then all classes (i) are moved down to the preceding class ($i-1$). The number of times that the classes are reset is an estimate of the number of beneficial sweeps that have passed through the population. Technically, these sweeps are not of particular alleles but of fitness classes. Thus we are measuring the rate at which better fitness classes replace worse ones, which is a reasonable measure of the rate of adaptation. A less direct estimate of the number of sweeps comes from the changing frequency of the marker allele [**34**]. The marker is initially absent from the population and increases in frequency according to its mutation rate. When beneficial mutations occur, they tend not to occur on marker chromosomes, which are fairly rare, but rather on non-marker chromosomes. Thus as beneficial sweeps occur, the non-marker chromosomes are at a selective advantage and tend to displace the marker chromosomes from the population. Eventually mutations to the marker begin to accumulate on the lineage with the new beneficial mutation and the marker once again rises in frequency. In keeping with the definition used in the experimental system [**34**], a sweep was identified when the marker frequency decreased over at least two censuses, where a census occurred approximately once every seven generations. The number of sweeps over a period of 10,000 generations was then determined. For comparison, Paquin and Adams observed 64 changes in marker frequency in 2,612 generations. Taking into account their observation that 1.6 times as many sweeps occurred in the diploid population, this observation translates into 188 sweeps (S_H, say) per 10,000 generations for the haploid populations and 302 sweeps (S_D) per 10,000 generations for the diploid populations.

To evaluate the importance of selection versus mutation on the rate of adaptation in the experiments of Paquin and Adams, simulations were performed using the data previously cited ($N_H = N_D = 5 \times 10^9$, $s = 0.1$, and $h = 0.9$) and allowing the remaining parameter, the rate of beneficial mutations, to vary. From equations (3.3) and (3.4), the diploid mutation rate should be $2h$ times as large as the haploid mutation rate, which allows for both the fact that twice as many mutations would appear in the diploid population and also the fact that proportionately fewer mutations survive loss when they first appear in diploids (since their fitness is smaller by an amount h). The haploid mutation rate was thus varied from 10^{-14} to 10^{-3} per individual ($= 2\mu Ls$ in terms of equation (3.3)); the diploid mutation rate was set equal to $2h$ times this quantity.

Only with the very smallest mutation rates ($\leq 10^{-13}$) did more mutations accumulate in the diploid population by an amount approaching $2h$. With such low mutation rates, mutations did not tend to cosegregate but rather accumulated independently. Note that from our analysis, we expect this independence when

the rate of fixation is faster than the rate of mutation; in accord with these simulations, $\rho_f \approx 0.002 \gg \rho_m$ implies that $2\mu Ls \ll 4 \times 10^{-13}$. When the mutation rate is this low, however, very few sweeps (< 10) occurred in 10,000 generations. Clearly this range of mutation rates is inappropriate for the observed number of sweeps ($S_H = 188$ and $S_D = 302$).

The number of sweeps observed for higher mutation rates is shown in Figures 6 and 7. In Figure 6, the number of times that the fitness classes were reset is shown, while, in Figure 7, the number of sweeps estimated by tracking the marker frequency is shown. For very high mutation rates (higher than 10^{-6}), the marker frequency was a very poor indicator of the number of beneficial sweeps; the marker frequency tended to increase and decrease very rapidly, with these fluctuations becoming uncoupled from the increase of any particular beneficial mutation. Over the entire range of parameters illustrated, haploids and diploids performed remarkably similarly. That is, even though diploid individuals experienced a mutation rate that was 80% larger than that of haploids, the fact that fitness was higher in haploids by only 1% was enough to compensate for this mutation difference. In short, in the range examined, small selective differences were enough to balance large differences in mutation rates, such that diploids did not adapt at a faster rate in the simulations.

Notice that, in Figure 7, the maximum number of sweeps observed in both the haploid and diploid populations occurred at mutation rates near 10^{-6}. This maximum corresponded to only 150 sweeps per 10,000 generations (as measured by changes in marker frequency), which is below the observed number of sweeps for both populations and far below the observed number for the diploid populations (again, $S_H = 188$ and $S_D = 302$). Nevertheless, taking 10^{-6} as the best estimate for the genome-wide rate of beneficial mutations, the simulations confirm that under the experimental conditions diploid populations would not have adapted faster than haploid populations simply because there were twice as many genes that could mutate beneficially. Further experiments are needed to shed light on this stark contradiction between the theoretical and experimental results and to determine, in other organisms, the rate of beneficial mutations and their selective consequences.

3.2.2. *Interbreeding haploid and diploid populations.* We now examine sweeps of beneficial mutations occurring within a sexual population and determine their influence on ploidy levels. We return to the life cycle illustrated in Figure 1, with genes at the ploidy locus altering the proportion of the life cycle spent in the haploid versus diploid phase. We now consider the impact of a single new beneficial mutation with selective coefficients given in Table 2. Let us first calculate the mean fitness among haploids and diploids. In any generation, if the mutation is at a frequency of p *among all genotypes at the ploidy locus*, then the mean fitness among individuals experiencing selection entirely as diploids would

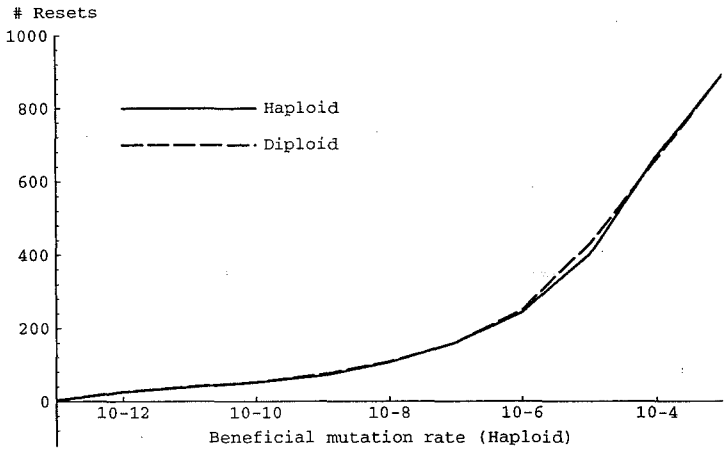

FIGURE 6. The number of sweeps observed in 10,000 generations, as evidenced by the disappearance of the wildtype class and the resetting of fitness classes. The selective advantage of mutations was 0.09 for diploids and 0.1 for haploids. The mutation rate appropriate to haploid individuals is given along the abscissa; the diploid mutation rate was obtained by multiplying this quantity by $2h$ (1.8).

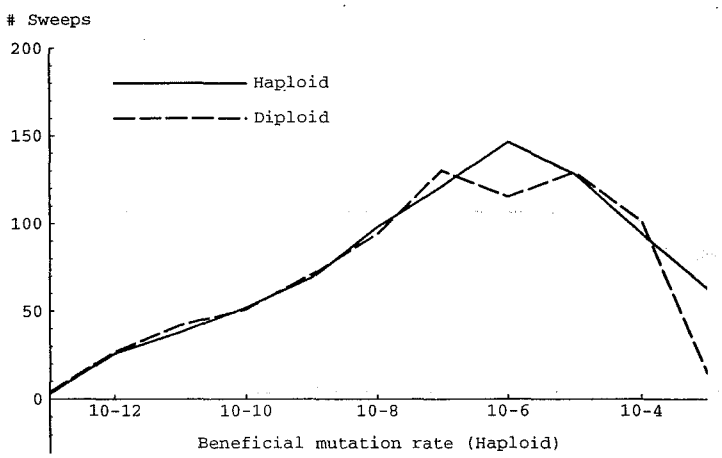

FIGURE 7. The number of sweeps observed in 10,000 generations, as evidenced by changes in the marker frequency. See caption for Figure 6.

be

$$\begin{aligned} \bar{w} &= (1+s)p^2 + (1+hs)2p(1-p) + (1-p)^2 \\ &= 1 + sp[p + 2h(1-p)], \end{aligned} \tag{3.10}$$

compared to

$$\bar{w} = (1+s)p + (1-p) = 1 + sp, \tag{3.11}$$

for individuals experiencing entirely haploid selection. It is easy to show that the mean fitness of diploids is higher than that of haploids when $h > 1/2$; conversely, the mean fitness of haploids is higher when $h < 1/2$. Even though a haploid that carries a beneficial mutation will tend to leave more offspring, a diploid has twice as many chances to carry a mutation; these factors balance at $h = 1/2$ such that haploids and diploids within a single interbreeding population have the same mean fitness (as long as the gene frequency is the same for both). For individuals spending a proportion (t) of the life cycle in the diploid phase and the remaining time $(1-t)$ in the haploid phase, the mean fitness is

$$\bar{w} = (1+s)p^2 + (1+hs)^t[(1+s)^{(1-t)} + 1]p(1-p) + (1-p)^2. \tag{3.12}$$

For weak selection (s near zero), the mean fitness is an increasing function of the amount of time spent in the diploid phase for dominant beneficial mutations ($h > 1/2$) but a decreasing function of t for recessive mutations ($h < 1/2$). For strong selection, the mean fitness is a more complex function of the parameters, but as a first approximation we expect the more diploid members of a population to have a higher average fitness and to increase in frequency during the sweep of new *dominant* beneficial mutations, while we expect the more haploid members of a population to increase in frequency during the sweep of new *recessive* mutations. These arguments, however, ignore the development of genetic associations, which may have a large impact on the evolution of ploidy levels, as was the case for deleterious mutations. It need not be the case that during the evolutionary process, the frequency of beneficial mutations is the same in the more haploid individuals and the more diploid individuals. Through simulations, we follow changes in ploidy levels that occur in response to sweeps of beneficial mutations. We find that genetic associations do develop that influence life cycle evolution.

The population evolves according to the recursions given in the Appendix A assuming positive selection and no recurrent mutations ($\mu_1 = \mu_2 = 0$). Initially, we will assume that the population is fixed upon a wildtype viability allele (A). In the absence of variation at the viability locus, evolution at the ploidy locus is neutral so that any combination of ploidy alleles is neutrally stable. The population under consideration thus begins with some positive frequency of both the C_1 and C_2 ploidy alleles ($0 < x_1, x_3 < 1; x_1 + x_3 = 1$). We then introduce a small amount of the beneficial mutation (a) and iterate the recursions until this new allele is near fixation within the population. At this point, we determine how the ploidy alleles have changed in frequency. A particular ploidy level is said to

be *favored*, when alleles that cause individuals to spend more time within this phase increase in frequency. Directional ploidy determination ($d_{11} < d_{12} < d_{22}$ or $d_{11} > d_{12} > d_{22}$) was assumed to simplify the interpretation of the results. We start by presenting results for the case when the new beneficial mutation is begun on both the C_1 and C_2 chromosomes according to their frequencies within the population (no initial disequilibrium). In this case, while the frequency of mutant *chromosomes* does not differ between haploids and diploids, the frequency of mutant *individuals* is higher in diploids, since diploid individuals have two alleles either of which may be mutants. We then present and discuss results for starting conditions with linkage disequilibrium. Disequilibrium would be expected if mutations tend to occur only once, initially with either the C_1 or the C_2 allele.

For any given rate of recombination and strength of selection, a value for the degree of dominance of the beneficial mutation (h) could be found above which increases in the diploid phase were favored and below which increases in the haploid phase were favored. This phenomenon is illustrated in Figure 8 for a particular starting position ($x_1 = 0.499$, $x_2 = 0.001$, $x_3 = 0.499$, $x_4 = 0.001$) and a particular set of ploidy alleles ($d_{11} = 0.499$, $d_{12} = 0.500$, $d_{22} = 0.501$ as in Figure 2). Twelve different combinations of ploidy alleles and starting positions (chosen to be different; all with no linkage disequilibrium) were similarly analyzed. Each produced slightly different cut-off values, but the differences were so slight as to make no qualitative and few quantitative differences to the graphs shown in Figure 8. These graphical analyses indicate that the evolution of diploidy is more often favored when beneficial mutations are dominant as expected, but that dominance was not a sufficient condition for the evolution of diploidy. In fact, with tight linkage or strong selection, haploidy was often favored despite the fact that the new beneficial mutations were dominant. This dependence on recombination implies that genetic associations do have a role in the evolution of ploidy levels during the spread of beneficial mutations in sexual populations.

Genetic associations were followed from one thousand random starting conditions (d_{ij} randomly chosen so as to be directional; x_i randomly chosen such that there was no linkage disequilibrium and so that $x_1 + x_3 \approx 0.99$) for each of five rates of recombination ($r \in 0, 0.01, 0.1, 0.25, 0.5$). Each run began with no disequilibrium and yet each run soon developed disequilibrium that favored the more haploid ploidy allele (negative disequilibrium if $d_{11} < d_{12} < d_{22}$ and positive disequilibrium if $d_{11} > d_{12} > d_{22}$). The simulations were then repeated with the linkage disequilibrium artificially reset to zero each generation using that generation's gene frequencies. By this method, we could remove the influence of disequilibrium on ploidy evolution for comparison. Under these conditions and for all rates of recombination, diploidy was favored when the beneficial mutations were dominant, while haploidy was favored when the beneficial mutations were recessive (the exact cut-off was slightly different from $h = 1/2$, but was consistent

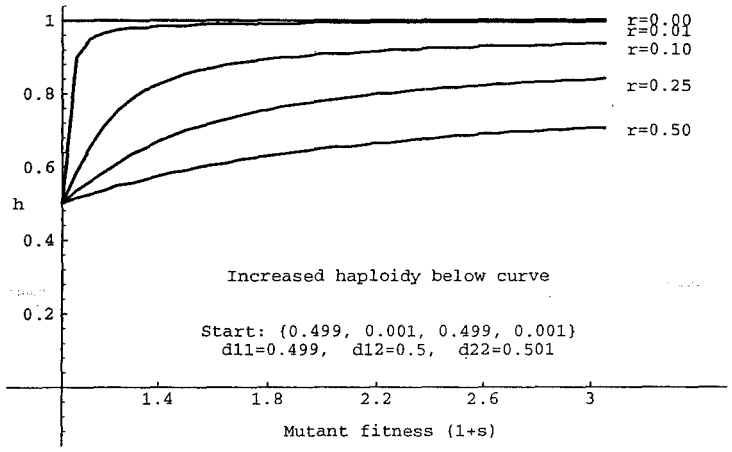

FIGURE 8. Parameters favoring the evolution of increased haploidy (below curve) or increased diploidy (above curve) for various rates of recombination (no initial disequilibrium). Nearly identical curves were produced with different starting positions (in linkage equilibrium) or with different ploidy alleles.

with equation (3.12)). The evolution of ploidy levels in response to beneficial mutations thus exhibits very similar behavior to the evolution of ploidy levels in response to deleterious mutations: ignoring disequilibrium, masking deleterious recessive mutations or revealing dominant beneficial mutations favors diploidy; genetic associations develop, however, that tend to favor the evolution of haploidy especially when linkage is tight.

A single sweep of a beneficial mutation can have a dramatic effect on the frequencies of the ploidy alleles within a population. In Table 4, the change in the frequency of a haploid ploidy allele is given for a range of parameter values ($h = 1/2$ so that haploidy is expected to be favored). A fairly rare haploid ploidy allele can quadruple in frequency, while more common ploidy alleles can traverse halfway across the frequency space towards fixation. These dramatic changes were observed when the rate of recombination was low and selection was weak. With low rates of recombination, weak selection is more effective in changing the frequencies of the ploidy alleles, while the opposite is true with high rates of recombination. The extent of frequency change appears to be related to the strength of genetic associations that are created and to the number of generations that these associations are maintained.

Heretofore, we have considered beneficial mutations that arise in a population in linkage equilibrium. The evolution of ploidy levels in response to sweeps of beneficial mutations is, however, sensitive to the initial level of linkage disequilibrium. With $d_{11} < d_{12} < d_{22}$, if the linkage disequilibrium is initially negative, then the beneficial mutation arises more frequently than expected on the haploid

	Initial freq(C_1)	Final freq(C_1) when $1+s$ equals: 1.20	1.50	2.00	10.0
$r = 0.0$	0.500	0.782	0.768	0.750	0.667
$r = 0.5$	0.500	0.509	0.520	0.532	0.566
$r = 0.0$	0.050	0.200	0.181	0.162	0.103
$r = 0.5$	0.050	0.052	0.054	0.057	0.065
$r = 0.0$	0.950	0.983	0.982	0.981	0.973
$r = 0.5$	0.950	0.952	0.954	0.956	0.961

TABLE 4. Extent of allele frequency change. The table shows the frequency of the C_1 allele (the more haploid allele) after a single sweep for three different starting frequencies of the C_1 allele (0.5, 0.05, and 0.95), two recombination rates (0 and 0.5), and four fitness values ($1+s = 1.20$, $1+s = 1.50$, $1+s = 2.00$, $1+s = 10.00$). The remaining parameters were $d_{11} = 0.2$, $d_{12} = 0.5$, $d_{22} = 0.8$, and $h = 0.5$. Under these conditions, haploidy is always favored as illustrated in Figure 8, but the extent of change depends on s and r.

allele (C_1) and haploidy is favored in a larger area of the parameter space. With positive disequilibrium, diploidy is favored for more combinations of parameters. In Figures 9 and 10, the curves delimiting the parameter space in which haploidy is favored from the space in which diploidy is favored are given for rather extreme initial disequilibrium values (only three of the four possible haplotypes exist within the starting population). When $r = 0$, the ploidy allele on which all beneficial mutations occur will sweep to fixation, regardless of its ploidy level. With higher rates of recombination, the selective regime will matter as well as the initial population composition. These considerations lead us to speculate that ploidy levels may fluctuate more often in populations with tight linkage (or other mechanisms that maintain disequilibria), since these fluctuations will simply reflect the random appearance of beneficial mutations on chromosomes with different ploidy alleles. With loose linkage or when beneficial mutations tend to appear in linkage equilibrium, the above results lead us to expect the evolution of haploidy when r is small, h is small, *or* s is large and to expect the evolution of diploidy when r is large, h is large, *and* s is small.

We clearly need more experimental evidence on the rate of beneficial mutations and their selective consequences. Only with this information can we determine whether sweeps of mutations tend to favor the evolution of diploidy in a particular organism. We know now, however, that beneficial mutations do not unequivocally favor the evolution of diploidy in sexual populations. Even though diploids have twice the opportunity to carry new mutations (as heterozygotes), the fact that these mutations can rise in frequency faster in haploids (especially when h is low) can overwhelm this advantage.

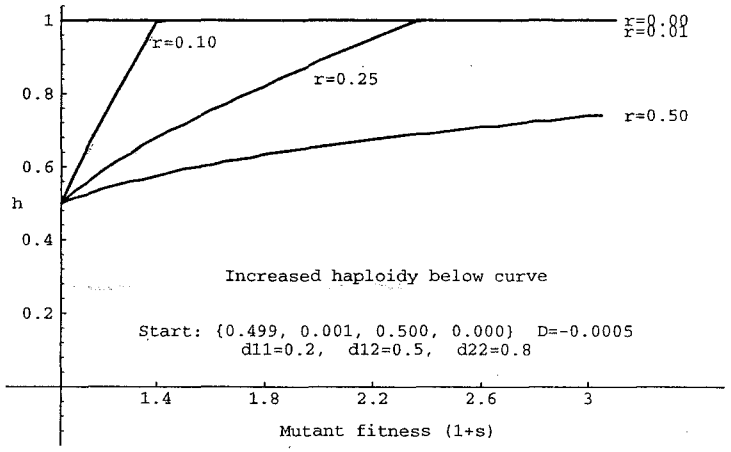

FIGURE 9. Parameters favoring the evolution of haploidy (below curve) or diploidy (above curve) for various recombination rates ($D = -0.0005$). Here $d_{11} = 0.2$, $d_{12} = 0.5$, and $d_{22} = 0.8$. If the d_{ij}s are chosen as in Figure 8 (closer together) then haploidy is favored even more often (haploidy favored for all h and r when $1 + s > 1.25$).

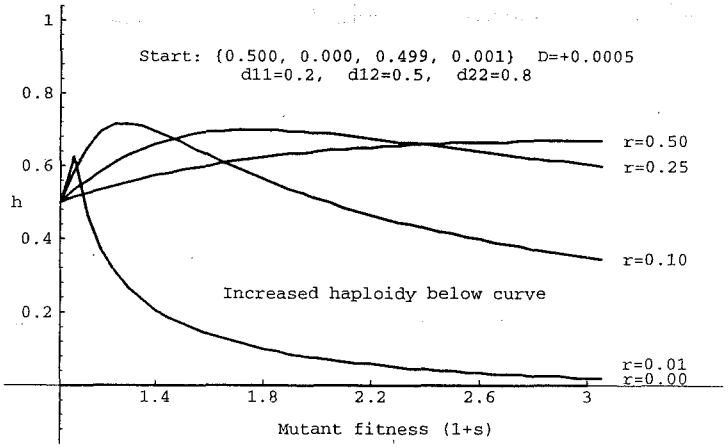

FIGURE 10. Parameters favoring the evolution of haploidy (below curve) or diploidy (above curve) for various recombination rates ($D = +0.0005$). Here $d_{11} = 0.2$, $d_{12} = 0.5$, and $d_{22} = 0.8$. If the d_{ij}s are chosen as in Figure 8 (closer together) then diploidy is favored even more often (diploidy favored for all h and r when $1 + s > 2.35$).

4. Conclusions

In a series of papers published during the 1920's [**17, 18**], Haldane showed that selection would occur most rapidly in populations that are either asexual or haploid. Having examined various forms of selection in amphimictic diploids (including autosomal, sex-linked, sex-limited and familial selection), Haldane concluded that "selection proceeds more slowly with all other systems of inheritance" [**17**]. The fact that allele frequencies change faster in response to selection acting on haploid rather than diploid genotypes in sexual organisms has not, until recently, been incorporated into theories about life cycle evolution. We now proceed to reiterate and interpret the main results in light of this fact.

It has generally been assumed that life cycles evolve as if allele frequencies were changing in a uniform manner throughout a population. Assuming that allele frequencies are equivalent among more haploid and more diploid members of a sexual population, equations (3.10) and (3.11) show that the mean fitness of diploids is higher than the mean fitness of haploids when deleterious mutations are recessive (masked) and when beneficial mutations are dominant (revealed). This leads one to expect that masking deleterious mutations and revealing beneficial mutations should favor the evolution of diploidy. This conclusion is erroneous, however, since allele frequencies will not remain uniform throughout a mixed population of haploids and diploids but will develop patterns that reflect the argument of Haldane. Specifically, allele frequencies will respond faster to selection among individuals which experience longer periods of haploid selection, with deleterious alleles decreasing in frequency and beneficial alleles increasing in frequency at faster rates. Put another way, since haploid individuals do not mask their mutations from selection, individuals that survive a long period of haploid selection tend to bear fewer offspring carrying deleterious alleles and more offspring carrying beneficial alleles than individuals that survive diploid selection. In other words, genetic associations tend to develop that couple alleles which increase the haploid phase with viability alleles that are the most fit.

Mechanisms that maintain genetic associations within a sexual population favor the expansion of the haploid phase in the life cycle. The most obvious such mechanism is tight linkage, but assortative mating, selfing, and asexual reproduction in populations with life cycle variation have similar influences on the evolution of ploidy levels [**33**]. For the increase of the diploid phase of a life cycle to be favored in an organism, mating must be fairly random within the population, recombination rates must be sufficiently high, and mutations must tend to be recessive when deleterious or dominant when beneficial.

Haldane's results also provide insight when we compare separate haploid and diploid asexual populations. Now selection is less effective in diploids not because of amphimixis, but because the effects of beneficial mutations are somewhat diluted by the presence of a second homologous allele within the genome

(assuming $h < 1$). Thus while diploids have twice as many opportunities to bear new beneficial mutations, these mutations will sweep to fixation at slower rates among diploid asexuals. The faster a beneficial mutation sweeps through an asexual population, the sooner new beneficial mutations can accumulate on lineages with previous beneficial mutations. Thus beneficial mutations can be nested within a lineage at a faster rate in haploid asexuals than diploid asexuals. This effect will be unimportant if mutations are rare and fix before the appearance of new mutations but will be very important if beneficial mutations occur often within a population. Even if mutations are rare, haploid populations will still accumulate beneficial mutations more rapidly than diploid populations if these mutations are recessive. This follows from other work by Haldane [**19**], which indicates that a beneficial mutation which arises in a single individual within a population will survive the initial period of random sampling with a probability that is approximately equal to the selective advantage of the new allele. This probability is $2s$ for new alleles within the haploid population but only $2hs$ in the diploid population. Twice as many mutations will still arise in a diploid population, but, if $h < 1/2$, fewer mutations will survive the first few generations of sampling when compared to a haploid population. The analyses in this paper allow us to evaluate whether mean fitness will increase at a higher rate in asexual haploid populations or in asexual diploid populations. Experiments are needed to determine whether haploid populations are favored under the conditions predicted (e.g. frequent mutations within the population) and are not favored when predicted (e.g. mutations rare within the population). Such experiments could manipulate the mutation rate within a population by either changing the population size or by exposing individuals to mutagens. Experiments are also needed to determine whether mean fitness differences measured in separate populations translate into competitive differences when the haploids and diploids are placed within a single environment.

The conditions outlined in this paper and elsewhere [**32, 33**] allow us to test the importance of masking deleterious mutations and revealing beneficial mutations to the evolution of ploidy levels. Given information about the life history of an organism and about selection, we can now predict whether the haploid or the diploid phase of the life cycle is likely to dominate. Whether these theories have any explanatory power has yet to be determined by careful comparative studies and experimental work.

5. Acknowledgements

The author wishes to thank Michael Cummings and Allen Orr for many fruitful discussions as well as Cheryl Jenkins, Mark Kirkpatrick, Maria Orive, and Monty Slatkin for innumerable suggestions that improved the presentation of this work.

Appendix A. Recursions.

In this appendix, we present and analyze recursions that monitor changes in the timing of meiosis in sexual organisms that experience both diploid and haploid selection (a regular alternation of generations). We census at the gamete stage of the life cycle, and then follow the four chromosomes (AC_1, aC_1, AC_2, aC_2) through the life cycle, tracking their frequencies over time (x_1, x_2, x_3, x_4, respectively). The recursions that map the population from one generation to the next and which correspond to the model described in the text and in Figure 1 are:

$$\begin{aligned}
Tx_1' &= (1-\tfrac{\mu_2}{2})(1-\tfrac{\mu_1}{2})[x_1x_1 + (1\pm hs)^{d_{11}}x_1x_2 + x_1x_3 \\
&\quad +(1\pm hs)^{d_{12}}(1-r)x_1x_4 + (1\pm hs)^{d_{12}}rx_2x_3]
\end{aligned}$$

$$\begin{aligned}
Tx_2' &= (\tfrac{\mu_2}{2})(1-\tfrac{\mu_1}{2})[x_1x_1 + (1\pm hs)^{d_{11}}x_1x_2 + x_1x_3 \\
&\quad +(1\pm hs)^{d_{12}}(1-r)x_1x_4 + (1\pm hs)^{d_{12}}rx_2x_3] \\
&\quad +(\tfrac{\mu_1}{2})[(1\pm s)^{(1-d_{11})}x_1x_1 + (1\pm hs)^{d_{11}}(1\pm s)^{(1-d_{11})}x_1x_2 \\
&\quad +(1\pm s)^{(1-d_{12})}x_1x_3 + (1\pm hs)^{d_{12}}(1\pm s)^{(1-d_{12})}(1-r)x_1x_4 \\
&\quad +(1\pm hs)^{d_{12}}(1\pm s)^{(1-d_{12})}rx_2x_3] \\
&\quad +[(1\pm hs)^{d_{11}}(1\pm s)^{(1-d_{11})}x_1x_2 + (1\pm hs)^{d_{12}}(1\pm s)^{(1-d_{12})}rx_1x_4 \\
&\quad +(1\pm s)x_2x_2 + (1\pm hs)^{d_{12}}(1\pm s)^{(1-d_{12})}(1-r)x_2x_3 \\
&\quad +(1\pm s)x_2x_4]
\end{aligned}$$

$$\begin{aligned}
Tx_3' &= (1-\tfrac{\mu_2}{2})(1-\tfrac{\mu_1}{2})[x_1x_3 + (1\pm hs)^{d_{12}}rx_1x_4 + (1\pm hs)^{d_{12}}(1-r)x_2x_3 \\
&\quad +x_3x_3 + (1\pm hs)^{d_{22}}x_3x_4]
\end{aligned}$$

$$\begin{aligned}
Tx_4' &= (\tfrac{\mu_2}{2})(1-\tfrac{\mu_1}{2})[x_1x_3 + (1\pm hs)^{d_{12}}rx_1x_4 + (1\pm hs)^{d_{12}}(1-r)x_2x_3 \\
&\quad +x_3x_3 + (1\pm hs)^{d_{22}}x_3x_4] \\
&\quad +(\tfrac{\mu_1}{2})[(1\pm s)^{(1-d_{12})}x_1x_3 + (1\pm hs)^{d_{12}}(1\pm s)^{(1-d_{12})}rx_1x_4 \\
&\quad +(1\pm hs)^{d_{12}}(1\pm s)^{(1-d_{12})}(1-r)x_2x_3 + (1\pm s)^{(1-d_{22})}x_3x_3 \\
&\quad +(1\pm hs)^{d_{22}}(1\pm s)^{(1-d_{22})}x_3x_4] \\
&\quad +[(1\pm hs)^{d_{12}}(1\pm s)^{(1-d_{12})}(1-r)x_1x_4 + (1\pm hs)^{d_{12}}(1\pm s)^{(1-d_{12})}rx_2x_3 \\
&\quad +(1\pm s)x_2x_4 + (1\pm hs)^{d_{22}}(1\pm s)^{(1-d_{22})}x_3x_4 + (1\pm s)x_4x_4]
\end{aligned}$$

where T is the sum of the right hand sides. In these recursions, $\mu_1/2$ measures the mutation rate during meiosis (when diploids produce haploid offspring) and $\mu_2/2$ measures the mutation rate during gametogenesis (when haploids produce gametes). Except to evaluate the role played by the position of mutations, we

assume that $\mu_1 = \mu_2 = \mu$. The $\pm$ sign should read as a positive sign when beneficial mutations are considered and as a negative sign when deleterious mutations are considered.

Appendix B. Analysis with recurrent deleterious mutations.

The equilibrium frequencies when the C_1 allele is fixed are:

$$\begin{aligned} \hat{x}_1 &= 1 - \hat{x}_2 \\ \hat{x}_2 &= \frac{\mu[1 + (1-s)^{(1-d_{11})}]}{2[1 - (1-s)^{(1-d_{11})}(1-hs)^{(1-d_{11})}]}. \end{aligned} \tag{B.1}$$

Invasion of the C_2 allele from near this equilibrium is governed by the leading eigenvalue obtained from the recursions linearized in the vicinity of the equilibrium (B.1) (see [**32**]). This leading eigenvalue equals:

$$\lambda_L = 1 - \mu K_1 + r\mu K_2 + O(\mu^2) \tag{B.2}$$

where

$$K_1 = \frac{[1 + (1-s)^{(1-d_{11})}][(1-hs)^{d_{11}} - (1-hs)^{d_{12}}]}{2[1 - (1-s)^{(1-d_{11})}(1-hs)^{d_{11}}]}$$

$$K_2 = \frac{(1-s)^{(1-d_{11}-d_{12})}(1-hs)^{d_{12}} K_3}{2[1 - (1-s)^{(1-d_{11})}(1-hs)^{d_{11}}][1 - (1-r)(1-s)^{(1-d_{12})}(1-hs)^{d_{12}}]}$$

$$\begin{aligned} K_3 = {} & (1-s)^{d_{11}} - (1-s)^{d_{12}} - (1-s)(1-hs)^{d_{11}} - (1-s)^{d_{12}}(1-hs)^{d_{11}} \\ & + (1-s)(1-hs)^{d_{12}} + (1-s)^{d_{11}}(1-hs)^{d_{12}}. \end{aligned}$$

K_1, K_2, and K_3 all have the same sign as $(d_{12} - d_{11})$. An inspection of the leading eigenvalue indicates that haploidy is favored whenever $r = 0$, but that with sufficiently strong masking and high recombination diploidy is favored (see equation (B.3)).

Quantitatively, the invasion of a new allele increasing the diploid phase of the life cycle will occur if

$$\begin{aligned} & rW_1^{d_{12}}[W_2^{d_{11}} - W_2^{d_{12}} - 2W_2W_1^{d_{11}} - W_2^{d_{11}}W_1^{d_{11}} \\ & \quad - W_2^{d_{12}}W_1^{d_{11}} + 2W_2W_1^{d_{12}} + 2W_2^{d_{11}}W_1^{d_{12}}] \\ & > W_2^{d_{12}}[1 + W_2^{(1-d_{11})}][W_1^{d_{11}} - W_1^{d_{12}}][1 - W_2^{(1-d_{12})}W_1^{d_{12}}], \end{aligned} \tag{B.3}$$

where $W_1 = (1 - hs)$ and $W_2 = (1 - s)$. Conversely, a new allele increasing the haploid phase of the life cycle will invade if this inequality is reversed.

Where in the life cycle mutations occur does not affect the qualitative behavior of the model with respect to linkage. It does, however, affect the quantitative results, especially when s is large. It can be shown that the evolution of diploidy is facilitated when mutations occur during meiosis (sporulation) as opposed to

gametogenesis. This effect may be understood by the fact that mutations which occur during sporulation are subjected to haploid selection before the diploids are formed, thus reducing the frequency of mutations carried by the diploids.

REFERENCES

1. ADAMS, J. AND HANSCHE, P. E. 1974. Population studies in microorganisms I. Evolution of diploidy in *Saccharomyces cerevisiae*. Genetics 76: 327–338.
2. ADAMS, J. AND OELLER, P. W. 1986. Structure of evolving populations of *Saccharomyces cerevisiae*: Adaptive changes are frequently associated with sequence alterations involving mobile elements belonging to the Ty family. Proc. Nat. Acad. Sci USA 83: 7124-7127.
3. BELL, G. 1982. The Masterpiece of Nature: The Evolution and Genetics of Sexuality. University of California Press, Berkeley.
4. CAVALIER-SMITH, T. 1978. Nuclear volume control by nucleoskeletal DNA, selection for cell volume and cell growth rate, and the solution of the DNA C-value paradox. J. Cell Sci. 34: 247–278.
5. CHARLESWORTH, B. 1983. Adaptive evolution in the laboratory. Nature 302: 479–480.
6. CROW, J. AND KIMURA, M. 1965. Evolution in sexual and asexual populations. Am Nat. 99: 439–450.
7. D'AMATO, F. 1977. Nuclear Cytology in Relation to Development. Cambridge University Press, Cambridge.
8. DARLINGTON, C. D. 1958. The Evolution of Genetic Systems. Basic Books, New York.
9. DESTOMBE, C., VALERO, M., VERNET, P., AND COUVET, D. 1989. What controls haploid-diploid ratio in the red alga, *Gracilaria verrucosa*? J. Evol. Biol. 2: 317–338.
10. EBERSOLD, W. T. 1967. *Chlamydomonas reinhardi:* Heterozygous diploid strains. Science 157: 447–449.
11. ESHEL, I. AND FELDMAN, M. W. 1982. On evolutionary genetic stability of the sex ratio. Theor. Pop. Biol. 21: 430–439.
12. EWENS, W. J. 1969. Population Genetics. Methuen & Co. LTD, London.
13. EWENS, W. J. 1979. Mathematical Population Genetics. Springer-Verlag, New York.
14. FOWELL, R. R. 1969. Life cycles in yeast. *In* A. H. Rose and J. S. Harrison (Eds.), The Yeasts (pp. 461–471). Academic Press, New York.
15. GOFF, L. J. AND COLEMAN, A. W. 1990. DNA: Microspectrofluorometric studies. *In* K. M. Cole and R. G. Sheath (Eds.), Biology of the Red Algae (pp. 43–71). Cambridge University Press, Cambridge.
16. GOLDSTEIN, D. 1992. Heterozygote advantage and the evolution of a dominant diploid phase. Genetics 132: 1195–1198.
17. HALDANE, J. B. S. 1924. A mathematical theory of natural and artificial selection, Part I. Trans. Camb. Phil. Soc. 23: 19–40.
18. HALDANE, J. B. S. 1926. A mathematical theory of natural and artificial selection, Part III. Proc. Camb. Phil. Soc. 23: 363–372.
19. HALDANE, J. B. S. 1927. A mathematical theory of natural and artificial selection, Part V: Selection and Mutation. Proc. Camb. Phil. Soc. 23: 838–844.
20. HALDANE, J. B. S. 1990. The Causes of Evolution. Princeton University Press, Princeton.
21. HOSHAW, R. W., WANG, J.-C., MCCOURT, R. M., AND HULL, H. M. 1985. Ploidal changes in clonal cultures of *Spirogyra communis* and implications for species definition. Amer. J. of Bot. 72: 1005–1011.
22. HUXLEY, J. 1964. Evolution: The Modern Synthesis. John Wiley & Sons, Inc., New York.
23. JENKINS, C. D. 1993. Selection and the evolution of genetic life cycles. Genetics 133: 401–410.
24. KIMURA, M. AND CROW, J. F. 1969. Natural selection and gene substitutions. Genet. Res. Camb. 13: 127–141.
25. KIMURA, M. AND OHTA, T. 1969. Average number of generations until fixation of a mutant gene in a finite population. Genetics 61: 763–771.

26. KIRKPATRICK, M. AND JENKINS, C. D. 1989. Genetic segregation and the maintenance of sexual reproduction. Nature 339: 300–301.
27. KONDRASHOV, A. AND CROW, J. 1991. Haploidy or diploidy: which is better? Nature 351: 314–315.
28. LOCK, R. H. 1906. Recent progress in the study of variation, heredity, and evolution. E. P. Dutton & Co., New York.
29. MAYR, E. 1982. The Growth of Biological Thought. The Belknap Press of Harvard University Press, Cambridge.
30. MULLER, H. J. 1932. Some genetic aspects of sex. Am. Nat. 66: 118–138.
31. ORR, H. A. AND OTTO, S. 1993. Does diploidy increase the rate of adaptation? Genetics: Submitted.
32. OTTO, S. AND GOLDSTEIN, D. 1992. Recombination and the evolution of diploidy. Genetics 131: 745–751.
33. OTTO, S. AND MARKS, J. 1994. Masking mutations from selection: A boon or a bane? Manuscript 000: 000.
34. PAQUIN, C. AND ADAMS, J. 1983a. Frequency of fixation of adaptive mutations is higher in evolving diploid than haploid yeast populations. Nature 302: 495–500.
35. PAQUIN, C. AND ADAMS, J. 1983b. Relative fitness can decrease in evolving asexual populations of *S. cerevisiae*. Nature 306: 368–371.
36. PERROT, V., RICHERD, S., AND VALERO, M. 1991. Transition from haploidy to diploidy. Nature 351: 315–317.
37. RAPER, J. R. AND FLEXER, A. S. 1970. The road to diploidy with emphasis on a detour. Symp. Soc. Gen. Microbiol. 20: 401–432.
38. SCHMALHAUSEN, I. I. 1949. Factors of Evolution: The Theory of Stabilizing Selection. The Blakiston Co., Philadelphia.
39. SIMMONS, M. J. AND CROW, J. F. 1977. Mutations affecting fitness in *Drosophila* populations. Ann. Rev. of Genetics 11: 49–78.
40. STRASBURGER, E. 1894. The periodic reduction of the number of the chromosomes in the life-history of living organisms. Ann. Bot. 8: 281–316.
41. SVEDELIUS, N. 1926. An evaluation of the structural evidence for genetic relationships in plants: Algae. Proc. Int. Cong. Pl. Sci. Ithaca 1: 457–471.
42. WEISS, R. L., KUKORA, J. R., AND ADAMS, J. 1975. The relationship between enzyme activity, cell geometry, and fitness in *Saccharomyces cerevisiae*. Proc. Nat. Acad. Sci USA 72: 794–798.
43. WIENER, P., FELDMAN, M. W., AND OTTO, S. P. 1992. On genetic segregation and the evolution of sex. Evolution 46: 775–782.

DEPARTMENT OF INTEGRATIVE BIOLOGY, UC BERKELEY, BERKELEY CA 94720
Current address: Department of Integrative Biology, UC Berkeley, Berkeley CA 94720
E-mail address: sarah@lolich.berkeley.edu