Somatic Mutation Favors the Evolution of Diploidy

H. Allen Orr

Department of Biology, University of Rochester, Rochester, New York 14627

Manuscript received July 10, 1994
Accepted for publication November 18, 1994

ABSTRACT

Explanations of diploidy have focused on advantages gained from masking deleterious mutations that are inherited. Recent theory has shown that these explanations are flawed. Indeed, we still lack any satisfactory explanation of diploidy in species that are asexual or that recombine only rarely. Here I consider a possibility first suggested by EFROIMSON in 1932, by MULLER in 1964 and by CROW and KIMURA in 1965: diploidy may provide protection against somatic, not inherited, mutations. I both compare the mean fitness of haploid and diploid populations that are asexual and investigate the invasion of "diploidy" alleles in sexual populations. When deleterious mutations are partially recessive and somatic mutation is sufficiently common, somatic mutation provides a clear advantage to diploidy in both asexual and sexual species.

> It would take us too far afield here to discuss the evolutionary merits of haploidy versus diploidy of the somatic tissue. H. J. MULLER (1964)

DESPITE its prevalence, the evolutionary advantage of diploidy remains unclear. The simplest, and most popular, theory suggested that diploids benefit from masking the effects of partially recessive mutations. Although intuitively appealing, theory shows that at equilibrium between mutation and selection, diploid populations actually suffer twice the mutation load of haploid populations, reflecting the twofold higher rate of mutation in diploids (CROW and KIMURA 1965). Masking does not, therefore, confer a long-term advantage on diploidy. A large number of alternative explanations of diploidy have been offered (PAQUIN and ADAMS 1983; KONDRASHOV and CROW 1991; PERROT *et al.* 1991; CHARLESWORTH 1991; BENGTSSON 1992; GOLDSTEIN 1992). None, however, is wholly convincing (OTTO and GOLDSTEIN 1992; VALERO *et al.* 1992; ORR and OTTO 1994).

Here I reconsider a possibility that was first hinted at by EFROIMSON (1932) and MULLER (1964) and discussed by CROW and KIMURA (1965). These authors suggested that diploidy might provide protection from the effects of somatic mutation; although haploids suffer the full brunt of such mutations, diploids might benefit from masking the effects of those uninherited mutations that are partially recessive. It seems worth reconsidering this possibility for two reasons. First, for reasons that are not clear, the somatic mutation hypothesis has been almost completely ignored in recent discussions of diploidy. Second, the somatic mutation hypothesis has not been considered mathematically.

Unfortunately, the effect of somatic mutation on diploid *vs.* haploid fitness may not be as straightforward as it might first seem. For one thing, diploids suffer a twofold higher rate of somatic, as well as germline, mutation. Second, because inherited mutations reach much higher frequencies in diploids than haploids, somatic mutation in diploids may often inadvertently "unmask" these usually hidden mutations. Last, and most important, because they must often occur late in development, somatic mutations may not have nearly as large an effect on fitness as inherited mutations. It may not, therefore, be so easy for somatic mutation to compensate for the greater load suffered by diploids from inherited mutations.

I consider these effects here. I first compare the mean fitnesses of haploid and diploid populations that are asexual. I then investigate the evolution of diploidy in sexual populations; in this case, I ask whether a rare mutant that increases the chance that an individual will pass through life as a diploid can invade. As we will see, when deleterious mutations are partially recessive and somatic mutation is sufficiently common, diploidy is favored in both asexuals and sexuals.

RESULTS

Fitness model for somatic mutations: The A allele (or class of alleles) is wild type and fit, whereas the a allele (or class of alleles) is deleterious. An A allele somatically mutates to a sometime during development with probability m. Because somatic mutations must often arise late in development, and so affect only a portion of adult tissues, they probably, on average, have milder fitness effects than inherited mutations. We can take this into account as follows: the fitness of a genotype that has suffered somatic mutation is given by

$$w_{som} = w - k\Delta w_u, \quad (1)$$

Author e-mail: haorr@darwin.biology.rochester.edu

TABLE 1

Fitness model used to incorporate the effects of somatic mutation

Before[a]	After[b]	Fitness[c]
A	a	$1 - ks$
AA	Aa	$1 - khs$
Aa	aa	$1 - hs - (1 - h)ks$

[a] Individual's genotype before somatic mutation (first entry is a haploid).

[b] Genotype (for at least a patch of tissue) after somatic mutation.

[c] Fitness of the resulting individual. Notice that if $k = 1$, the somatic mutation has as severe an effect as an equivalent inherited mutation.

where w gives the genotype's fitness before somatic mutation and Δw_u is the amount by which fitness would be reduced *if* the somatic mutation had been inherited. Thus, k gives the average fitness effect of a somatic mutation relative to an equivalent inherited one ($0 \leq k \leq 1$). When $k = 1$, somatic mutations are as deleterious as inherited ones, whereas when $k = 0$, somatic mutations have no effect on fitness. In reality, the average fitness effects of somatic mutations might be quite small. Table 1 shows how the fitness of haploid and diploid genotypes at the A locus is decreased by somatic mutation.

In the calculations that follow, the terms m and k always enter as a product. Thus, it will prove useful to define an "effective" somatic mutation rate, $m_e = mk$. In essence, we discount the actual somatic mutation rate to reflect the fact that somatic mutations have smaller effects than inherited ones. An effective somatic rate of m_e reduces fitness by the same amount as inherited mutations occurring at the same rate.

Asexual species: We first consider the simple case where haploids and diploids do not interbreed. The cycle of events is formation of a zygote (*i.e.*, inheritance of germline mutations), somatic mutation, selection and reproduction.

In a haploid population, an inherited deleterious mutation reaches an equilibrium frequency of $\hat{q} = u/s$, where u is the germline mutation rate and s is the selection coefficient (CROW 1970). (As shown below, this classical mutation-selection equilibrium is changed somewhat by somatic mutation, but the effect is usually negligibly small.) Incorporating the effects of somatic mutation (Table 1), the mean fitness of haploids is just

$$\begin{aligned}\bar{w}_{hap} &= (1 - \hat{q})[(1 - m) + m(1 - ks)] + \hat{q}(1 - s) \\ &= 1 - u - mk(s - u) \\ &\approx 1 - u - m_e s, \end{aligned} \tag{2}$$

where the approximation ignores the product of germline and somatic mutation rates. Not surprisingly, somatic mutation reduces mean fitness by about $m_e s$ below that due to germline mutation.

In diploids, a deleterious mutation reaches an equilibrium frequency of $\hat{q} \approx u/hs$ where h is the dominance of the mutation ($h = 0$ means that the mutation is completely recessive and $h = 1$ that it is completely dominant). The approximation assumes that the mutation has some heterozygous effect, as shown by data from Drosophila (CROW 1970). Mutant homozygotes will be negligibly rare and thus, before somatic mutation occurs, mutant heterozygotes make up about $2q$ of the population. With somatic mutation, the mean fitness of diploids is

$$\begin{aligned}\bar{w}_{dip} &\approx (1 - 2\hat{q})[(1 - 2m) + 2m(1 - khs)] \\ &\quad + 2\hat{q}[(1 - m)(1 - hs) \\ &\quad + m(1 - hs - (1 - h)ks)] \\ &\approx 1 - 2u - 2mk\left(hs - 3u + \frac{u}{h}\right) \\ &\approx 1 - 2u - 2m_e hs. \end{aligned} \tag{3}$$

Thus, somatic mutation reduces the mean fitness of diploids by about $2m_e hs$ below that otherwise expected.

Therefore, $\bar{w}_{dip} - \bar{w}_{hap} \approx m_e s(1 - 2h) - u$, and diploids enjoy a higher mean fitness than haploids when

$$h < \frac{1}{2}\left(1 - \frac{u}{m_e s}\right). \tag{4}$$

The term in brackets shows that somatic mutation does not automatically favor diploidy whenever mutations are partially recessive ($h < \frac{1}{2}$). Rather, diploidy is favored only when somatic mutations are partially recessive *and* are common enough to offset the greater load suffered by diploids from inherited mutations. In particular, diploidy is favored only when the effective rate of somatic mutation exceeds

$$m_e > \frac{u}{s(1 - 2h)}, \tag{5}$$

where we assume $h < \frac{1}{2}$.

Interestingly, the rate of somatic mutation required to favor diploidy decreases with larger s. This is because—although the fitness difference between diploids and haploids due to inherited mutation is independent of s (HALDANE 1937)—the fitness difference due to somatic mutation is directly proportional to s. In the extreme, but biologically important, case of a nearly recessive lethal mutation (small h and $s = 1$), diploidy is favored whenever $m_e \gtrsim u$.

It is trivial to extend this model to multiple loci if different mutations have independent effects on fitness. With multiplicative fitness effects across loci, $\bar{w}_{hap} \approx \exp[-(U + M_e s)]$ and $\bar{w}_{dip} \approx \exp[-2(U + M_e hs)]$, where U and M_e are the germline and effective somatic mutation rates per haploid *genome*, respectively. Dip-

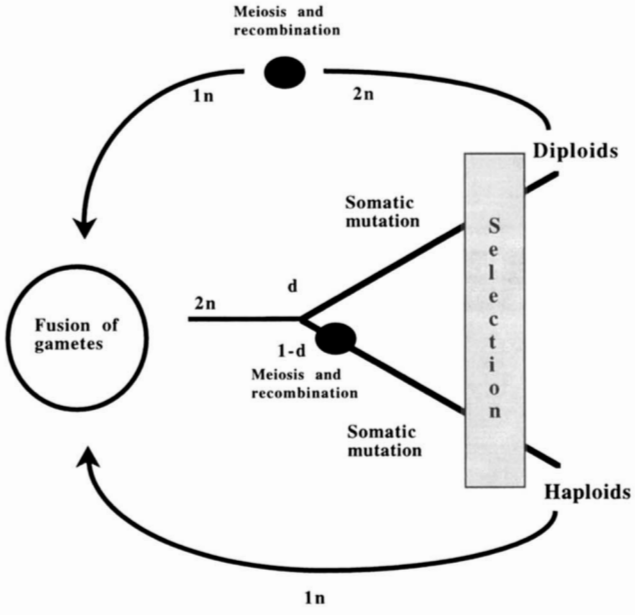

FIGURE 1.—Life cycle of the sexual species modeled (adapted from OTTO and GOLDSTEIN 1992). Haploid and diploid adults produce gametes at the same time, which randomly fuse. Germline mutations arise in gametes. Zygotes either remain diploid or immediately undergo meiosis and recombination to produce haploid adults. Somatic mutation occurs during haploid or diploid development and then selection occurs. Haploids produce gametes mitotically, whereas diploids produce gametes meiotically. A deleterious allele, a, is at mutation-selection equilibrium before the appearance of a new mutation (C_2) at the ploidy locus. This mutation affects the probability that an individual will be haploid *vs.* diploid at the time of selection. [This model yields results that are nearly identical to those of a more complex model that explicitly considers alternation between haploid and diploid phases (see OTTO 1994).] When $d_{11} = 0$ (C_1C_1 always haploid) and $d_{12} = 1$ (C_1C_2 always diploid), the model is equivalent to that of PERROT *et al.* (1991).

loids are favored when $M_e > U/[s(1 - 2h)]$ and $h < \frac{1}{2}$.

Somatic mutation in sexuals: The mean fitness arguments made above assume that individuals of different ploidy levels do not interbreed. Here I consider the more complex, but perhaps more biologically relevant, case in which haploids and diploids interbreed. In particular, I investigate the fate of a rare mutation that changes the probability that a zygote delays meiosis and remains diploid, when both germline and somatic mutation occur.

The model is an extension of those of PERROT *et al.* (1991) and OTTO and GOLDSTEIN (1992). Ploidy is determined by alleles at a single locus "C," whereas selection acts on some second locus "A" that affects adult viability. A deleterious allele, a, is produced at this locus by both germline and somatic mutation. The two loci recombine at a rate r. "Diploid" and "haploid" adults produce gametes at the same time, which then randomly fuse to produce zygotes (see Figure 1 for life cycle). Following OTTO and GOLDSTEIN, wild-type C_1C_1 zygotes remain diploid with probability d_{11} or become haploid with probability $(1\text{-}d_{11})$. A new mutation (C_2) arises at the ploidy locus. Heterozygotes for this mutation (C_1C_2) become diploid with probability d_{12} or haploid with probability $(1 - d_{12})$. The appearance of a new "diploidy" allele means that $d_{12} > d_{11}$. It should be emphasized that the ploidy locus itself is not under direct selection; alleles at the ploidy locus change frequency only in response to selection at the viability locus.

Derivation of exact recursion equations for the four gamete types (C_1A, C_1a, C_2A, C_2a) that take into account both germline and somatic mutation is tedious but straightforward. The results are shown in the APPENDIX. Once again, the terms m and k always enter as a product ($= m_e$). As expected, when there is no somatic mutation ($m = 0$) or when such mutations have no effect on fitness ($k = 0$), these recursions collapse to those presented by OTTO and GOLDSTEIN, who ignored somatic mutation.

From the recursions, one can show that before the appearance of the diploidy allele C_2, a deleterious allele segregates at an equilibrium frequency of

$$\hat{q} \approx \frac{u}{s[1 - d_{11}(1 - h) - m_e(1 - 2d_{11} + 3d_{11}h)]}. \quad (6)$$

When the population is nearly completely haploid ($d_{11} \approx 0$), this is just $\hat{q} \approx u/[s(1 - m_e)]$. In short, somatic mutation slightly increases the equilibrium frequency of *inherited* mutations. The reason is that by converting some A individuals to a, somatic mutation reduces the fitness difference between wild-type and mutant individuals. The APPENDIX discusses other aspects of mutation-selection balance with somatic mutation. For present purposes, we are interested in the fate of a diploidy allele that arises in a population near the equilibrium given by Equation 6.

I performed a local stability analysis to find the conditions under which a rare diploidy allele invades a haploid population. The details are given in the APPENDIX. To simplify our treatment (and to allow comparison with the asexuals studied above), we can consider the appearance of a complete diploid allele ($d_{12} = 1$) in a haploid population ($d_{11} = 0$). In this case, the leading eigenvalue given in the APPENDIX (Equation A5) shows that diploids will invade when

$$m_e > \frac{u}{s}\left[\frac{(1 - h)\,hs}{(1 - 2h)[r + (1 - r)\,hs]} - 1\right] \quad (7)$$

and $h < \frac{1}{2}$.

Equation 7 shows that *no matter what the rate of recombination, there is always some rate of somatic mutation above which diploidy is favored*. This is true even in the extreme case of no recombination: although OTTO and

GOLDSTEIN (1992) found that diploidy is never favored in nonrecombining species, Equation 7 shows that diploids *do* invade such species as long as

$$m_e > \frac{uh}{s(1 - 2h)} . \quad (8)$$

This condition is actually less stringent than that for asexuals (compare Equations 5 and 8). This is presumably because—when diploids and haploids interbreed—deleterious alleles do not rise to as high an equilibrium frequency as in an isolated diploid population. "Interbreeding" diploids do not, therefore, suffer a full diploid load of $2u$. Equation 8 also shows that in the biologically important case of a nearly recessive lethal mutation, diploidy is favored when $m_e \gtrsim uh$, *i.e.*, the somatic mutation rate need only be of the order of the germline mutation rate (or smaller).

The effect of somatic mutation on the evolution of diploidy is best seen graphically. Figure 2 shows the level of dominance, h, for deleterious alleles below which diploids will invade. We again focus on the worst-case scenario of rare recombination (in the top panel, $r = 0.01$, whereas in the bottom panel, $r = 0.1$). In the absence of somatic mutation ($m_e = 0$), diploidy alleles do not usually invade (OTTO and GOLDSTEIN 1992). Somatic mutation dramatically increases the parameter space over which diploids invade. The effect is quite large when the effective rate of somatic mutation is on the order of the germline mutation rate. If m_e is considerably larger than u, it is clear that diploidy alleles usually invade whenever deleterious mutations are partially recessive. As expected intuitively, diploidy is never favored when deleterious mutations are partially dominant ($h > \frac{1}{2}$), as diploidy cannot mask the effects of dominant mutations.

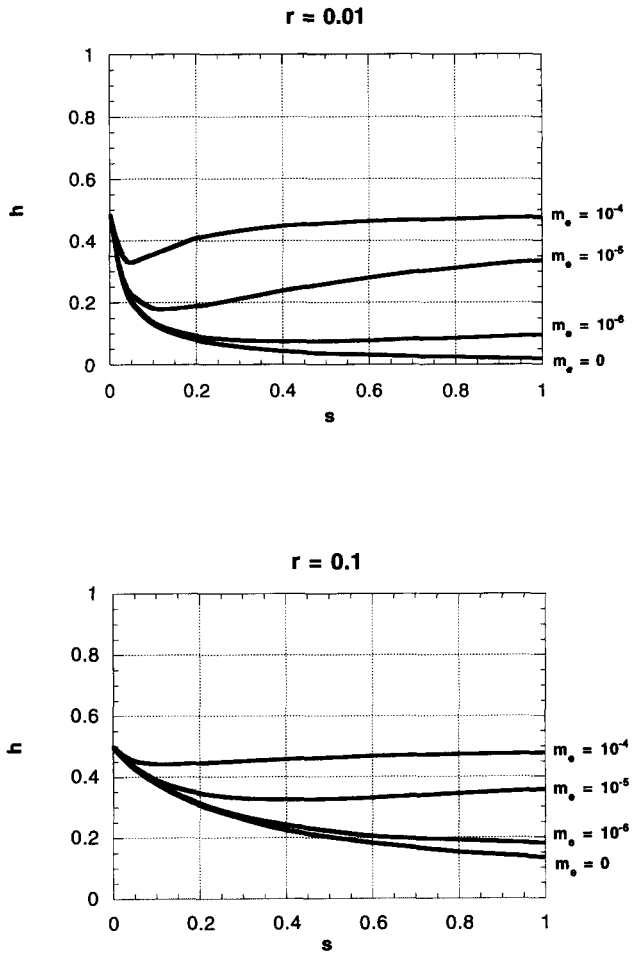

FIGURE 2.—The critical level of dominance, h, for deleterious mutations below which a "diploidy" allele invades. The curves correspond to different effective rates of somatic mutation ($m_e = mk$). In all cases, the germline mutation rate was $u = 10^{-5}$. Here we consider the case where the wild-type (C_1) allele at the ploidy locus was predominantly haploid ($d_{11} = 0.01$), whereas the new mutant allele (C_2) was predominantly diploid ($d_{22} = 0.99$, $d_{12} = 0.50$; the position of the curves is quite insensitive to the exact values of d_{ij}). Alleles at the viability locus (A) began at mutation-selection balance frequencies (see Equation 6) before the introduction of C_2. The diploidy mutant began at a very low frequency (=0.001) and appeared in linkage equilibrium with the viability locus. As expected, the $m_e = 0$ curves are identical to those in OTTO and GOLDSTEIN (1992).

DISCUSSION

Masking of partially recessive mutations usually confers no advantage on diploid populations. The reason is that although fairly recessive mutations have smaller fitness effects (hs) in diploids, they rise to higher equilibrium frequencies in diploid than haploid populations (heterozygotes occur at a frequency of $\approx 2u/hs$). In the end, the mutation load is independent of the dominance of inherited mutations ($L_{dip} \approx hs \times 2u/hs \approx 2u$) and diploids are no better off than haploids. In fact, diploids are actually worse off as they suffer from a twofold greater rate of mutation.

Masking of somatic mutations, on the other hand, *can* favor diploidy. The reason is obvious: populations cannot reach mutation-selection balance for mutations that are not inherited. Recessive somatic mutations do not somehow reach higher frequencies than more dominant ones, and so the load due to somatic mutation must depend on dominance (in an asexual, the load due to somatic mutation is $L_{dip} \approx 2m_e hs$, while $L_{hap} \approx m_e s$). Diploids can, therefore, enjoy a higher mean fitness than haploids if deleterious mutations are partially recessive and somatic mutation is common enough (Equation 5). Summed across loci, this fitness difference may be appreciable. Although these mean fitness arguments are essentially group selectionist, they are supported by study of the conditions under which diploidy alleles invade a haploid population: somatic mutation dramatically eases the conditions for invasion (see Figure 2 and the APPENDIX). It is especially important to note that this advantage to diploidy appears even when recombination is rare or absent; as OTTO and

GOLDSTEIN (1992) and ORR and OTTO (1994) emphasized, no previous theory offers a clear advantage to diploidy in asexuals or in sexuals who rarely recombine.

Given that CROW and KIMURA (1965) suggested long ago that protection from somatic mutation may confer an advantage to diploidy—and given that the calculations confirming this argument are straightforward—it is not clear why this possibility has been so thoroughly ignored in recent discussions of diploidy. One possibility is that because somatic mutations must often arise late in development (and so presumably have smaller effects on fitness), they may not seem a potent enough force to affect the evolution of ploidy level. The net effect of somatic mutation is, however, a function of both the mean effect of somatic relative to inherited mutations (k) *and* the probability that a somatic mutation arises during development (m). Although k obviously declines as somatic development becomes more complex (as a larger fraction of somatic mutations appear later in development), m must simultaneously increase (as the soma is becoming more complex).

It is easy to show that this latter effect predominates if the somatic mutation rate is a function of the number of cell divisions, as seems likely (*e.g.*, KUICK *et al.* 1992). Consider again a simple model with a single bout of selection at the end of somatic development but before reproduction. "M" now gives the *per cell division* rate of somatic mutation for a haploid genome. At time $t = 0$, there is only one cell; at time $t = 1$, there are two daughter cells, and so on. We let later somatic mutations have smaller effects than earlier ones. In particular, fitness effect is proportional to the proportion of adult tissue ultimately "inheriting" a somatic mutation; a somatic mutation that occurs in the single-cell stage ($t = 0$) has as large an effect as an inherited mutation, whereas one that occurs in the four-cell stage has, on average, a fourth of this effect, and so on. (This approximation is clearly invalid for mutations that affect cell viability *per se* or for mutations causing uncontrolled cell division. As a first approximation, however, it must be near enough true that the fitness effects of somatic mutations are inversely related to the time during development in which they arise.)

If there are t_r cell cycles before reproduction, a total of $2^{t_r} - 1$ cell divisions occur in development. Thus, we expect about $2^{t_r} M$ somatic mutations during haploid development (we assume each occurs at a unique site). Taking into account the effects of mutations that occur at different times in development, it is easy to show that the average somatic mutation has a fitness effect of $k \approx t_r/2^{t_r}$ $(= \Sigma_{t=0}^{t_r-1}[(2^t M)(2^{-t})]/2^{t_r}M)$. k obviously declines rapidly as somatic development becomes more complex (*i.e.*, as t_r increases). Nonetheless, the *total* fitness effect of somatic mutation is given by the product of ks and the total number of mutations, or

$$L \approx t_r M s. \quad (9)$$

In short, although the mean effect of a somatic mutation declines in more and more complex species, the total fitness effect of somatic mutation grows approximately linearly with the "length", t_r, of somatic development. Somatic mutation could, therefore, be a potent force in organisms, like metazoans, that possess a very complex soma.

Figure 2 shows that even with very low recombination rates, somatic mutation has a large effect on the evolution of diploidy when m_e is on the order of the germline mutation rate. The critical question then is how common is somatic mutation? There have been remarkably few attempts to directly measure somatic mutation rates. Perhaps the best data come from humans, because of the role of somatic mutation in many forms of cancer (*e.g.*, VOGELSTEIN *et al.* 1988). Several studies have estimated somatic mutation rates in human cells grown in culture (reviewed in LOEB 1991; KUICK *et al.* 1992). These studies have all yielded values near 10^{-6}–10^{-7} mutations per locus per cell division. Taking into account the sizes of the genes studied, the mean somatic mutation rate is near 1.4×10^{-10} mutations/nucleotide/cell division (LOEB 1991). Although this value is a rough estimate with many possible sources of error, it is remarkably close to the estimated rate of *germline* mutation in humans ($=1.2 \times 10^{-10}$ mutations/nucleotide/cell division) (LOEB 1991; KUICK *et al.* 1992).

The fact that $U \approx M$ (where U is the per division rate of genomic mutation) allows a simple comparison of the fitness effects of germline and somatic mutations. Because the effective rate of somatic mutation is $t_r M$ (from Equation 9), whereas the analogous germline mutation rate is $t_g U$ (where t_g is the number of germline cell "generations"), somatic mutation may have as large an effect on fitness as inherited mutation whenever there are more cell generations in somatic than germline development (*i.e.*, whenever $t_r > t_g$).

In any case, it is clear that an enormous number of somatic mutations must arise in species having an extended development. In humans, for instance, there may be about 10^{16} cell divisions throughout a normal life; taking into account the somatic mutation rate and the size of the human genome (coding regions only), the average human may suffer from 10^{13} somatic mutations (LOEB 1991). Although we do not know the average fitness effects of these mutants relative to the equivalent inherited mutations (indeed many must occur after peak ages of reproduction), their sheer number strongly suggests that somatic mutation cannot have negligible fitness effects. (A conclusion that is graphically reinforced by the high incidence of human cancers involving somatic mutation.)

In summary, somatic mutation favors the evolution of diploidy. Indeed, the somatic mutation hypothesis

seems an especially appealing explanation of diploidy for two reasons. First, somatic mutation favors diploidy in both sexual and asexual species. Second, and more important, the somatic mutation hypothesis makes sense of the broad phylogenetic distribution of haploidy *vs.* diploidy: although exceptions can be found, extended haploid phases are clearly far more common among developmentally simple organisms (*e.g.*, chlorophyta), whereas extended diploid phases are nearly universal among developmentally complex organisms (*e.g.*, metazoans and seed plants) (BELL 1982). It is not so obvious why this association should exist under the germline masking hypothesis (PERROT *et al.* 1991), the synergistic interaction hypothesis (KONDRASHOV and CROW 1991) or the heterosis hypothesis (GOLDSTEIN 1992). This pattern is, however, expected if the advantage of diploidy derives from protection against somatic mutation.

I thank B. BENGTSSON, B. CHARLESWORTH, A. CLARK, J. CROW, J. JAENIKE, A. KONDRASHOV, L. ORR, S. OTTO and an anonymous reviewer for helpful comments or discussion. This work was supported by funds from the Dobzhansky Prize.

LITERATURE CITED

BELL, G., 1982 *The Masterpiece of Nature: The Evolution and Genetics of Sexuality*. University of California Press, Berkeley.

BENGTSSON, B., 1992 Deleterious mutations and the origin of the meiotic ploidy cycle. Genetics **131:** 741–744.

CHARLESWORTH, B., 1991 When to be diploid. Nature **351:** 273–274.

CROW, J. F., 1970 Genetics loads and the cost of natural selection, pp. 128–177 in *Mathematical Topics in Population Genetics*, edited by J.-I. KOJIMA. Springer-Verlag, New York.

CROW, J. F., and M. KIMURA, 1965 Evolution in sexual and asexual populations. Am. Nat. **99:** 439–450.

EFROIMSON, V. P., 1932 On some problems of accumulation and action of lethals. Biol. J. **1:** 87–101 (in Russian).

GOLDSTEIN, D. B., 1992 Heterozygote advantage and the evolution of a dominant diploid phase. Genetics **132:** 1195–1198.

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

KONDRASHOV, A. S., and J. CROW, 1991 Haploidy or diploidy: which is better? Nature **351:** 314–315.

KUICK, R. D., J. V. NEEL, J. R. STRAHLER, E. H. Y. CHU, R. BARGAL *et al.*, 1992 Similarity of spontaneous germinal and *in vitro* somatic cell mutation rates in humans: implications for carcinogenesis and for the role of exogenous factors in "spontaneous" germinal mutagenesis. Proc. Natl. Acad. Sci. USA **89:** 7036–7040.

LOEB, L. 1991 Mutator phenotype may be required for multistage carcinogenesis. Cancer Res. **51:** 3075–3079.

MULLER, H. J., 1964 The relation of recombination to mutational advance. Mutat. Res. **1:** 2–9.

ORR, H. A., and S. P. OTTO, 1994 Does diploidy increase the rate of adaptation? Genetics **136:** 1475–1480.

OTTO, S. P., 1994 The role of deleterious and beneficial mutations in the evolution of ploidy levels, pp. 69–96 in *Theories for the Evolution of Haploid-diploid Life Cycles*, edited by M. KIRKPATRICK. American Mathematical Society, New York.

OTTO, S. P., and D. B. GOLDSTEIN, 1992 Recombination and the evolution of diploidy. Genetics **131:** 745–751.

PAQUIN, C., and J. ADAMS, 1983 Frequency of fixation of adaptive mutations is higher in evolving diploid than haploid yeast populations. Nature **302:** 495–500.

PERROT, V., S. RICHERD and M. VALERO, 1991 Transition from haploidy to diploidy. Nature **351:** 315–317.

VALERO, M., S. RICHERD, V. PERROT and C. DESTOMBE, 1992 Evolution of alternation of haploid and diploid phases in life cycles. Trends Ecol. Evol. **7:** 25–29.

VOGELSTEIN, B., E. R. FEARON, S. R. HAMILTON, S. E. KERN, A. C. PREISINGER *et al.*, 1988 Genetic alterations during colorectal-tumor development. N. Engl. J. Med. **319:** 525–532.

Communicating editor: A. G. CLARK

APPENDIX

Recursions with somatic mutation: The recursions presented below incorporate dominance, recombination, germline and somatic mutation rates for both haploid and diploid selection. Not surprisingly, they are rather complicated. In the absence of somatic mutation ($m_e = 0$), they reduce to those presented by OTTO and GOLDSTEIN (1992). In the simplest case of $d_{11} = 0$ and $d_{12} = 1$, we are considering the invasion of an allele that is always diploid into a population that is entirely haploid. This is equivalent to the situation considered by PERROT *et al.* (1991).

We define the following gamete frequencies:

$$x_1 = \text{frequency of } C_1A$$
$$x_2 = \text{frequency of } C_1a$$
$$x_3 = \text{frequency of } C_2A$$
$$x_4 = \text{frequency of } C_2a.$$

Taking into account all possible matings, the recursions for these gamete frequencies are

$$\begin{aligned}
x_1'\bar{w} = {} & (1-u)\{x_1^2 w_{AA,A\cdot d_{11}} + x_1x_2 w_{Aa,A\cdot d_{11}} \\
& + x_1x_3 w_{AA,A\cdot d_{12}} + (1-r)x_1x_4 w_{Aa,A\cdot d_{12}} \\
& + rx_2x_3 w_{Aa,A\cdot d_{12}}\} \\
x_2'\bar{w} = {} & u\{x_1^2 w_{AA,A\cdot d_{11}} + x_1x_2 w_{Aa,A\cdot d_{11}} + x_1x_3 w_{AA,A\cdot d_{12}} \\
& + (1-r)x_1x_4 w_{Aa,A\cdot d_{12}} + rx_2x_3 w_{Aa,A\cdot d_{12}}\} \\
& + \{x_1x_2 w_{Aa,a\cdot d_{11}} + rx_1x_4 w_{Aa,a\cdot d_{12}} + x_2^2 w_{aa,a\cdot d_{11}} \\
& + (1-r)x_2x_3 w_{Aa,a\cdot d_{12}} + x_2x_4 w_{aa,a\cdot d_{12}}\} \\
x_3'\bar{w} = {} & (1-u)\{x_1x_3 w_{AA,A\cdot d_{12}} + rx_1x_4 w_{Aa,A\cdot d_{12}} \\
& + (1-r)x_2x_3 w_{Aa,A\cdot d_{12}} \\
& + x_3^2 w_{AA,A\cdot d_{22}} + x_3x_4 w_{Aa,A\cdot d_{22}}\} \\
x_4'\bar{w} = {} & u\{x_1x_3 w_{AA,A\cdot d_{12}} + rx_1x_4 w_{Aa,A\cdot d_{12}} \\
& + (1-r)x_2x_3 w_{Aa,A\cdot d_{12}} + x_3^2 w_{AA,A\cdot d_{22}} \\
& + x_3x_4 w_{Aa,A\cdot d_{22}}\} + \{(1-r)x_1x_4 w_{Aa,a\cdot d_{12}} \\
& + rx_2x_3 w_{Aa,a\cdot d_{12}} + x_2x_4 w_{aa,a\cdot d_{12}} \\
& + x_3x_4 w_{Aa,a\cdot d_{22}} + x_4^2 w_{aa,a\cdot d_{22}}\}. \qquad (2)
\end{aligned}$$

The mean fitness, $\bar{w}$, is given by the sum of all the right-hand sides. The w terms are equal to

$$w_{AA,A\cdot d_{ij}} = 1 - m_e s[1 - d_{ij}(1 - 2h)]$$

$$w_{Aa,A\cdot d_{ij}} = 1 - s[m_e + (1 - m_e) d_{ij} h]$$

$$w_{Aa,a\cdot d_{ij}} = 1 - s[1 - d_{ij}(1 - h - m_e + m_e h)]$$

$$w_{aa,a\cdot d_{ij}} = 1 - s.$$

Mutation-selection balance with somatic mutation: Before the appearance of the C_2 allele at the ploidy locus, an inherited deleterious allele a reaches a mutation-selection equilibrium frequency given by Equation 6 in the text (this approximation ignores terms involving the products $m_e u$, $x_2 u$ and terms of x_2^2 and higher). As noted, if the population is entirely haploid ($d_{11} = 0$), a obtains a frequency of

$$\hat{q} = x_2 \approx \frac{u}{s(1 - m_e)}. \tag{A1}$$

In an entirely diploid population ($d_{11} = 1$), on the other hand, a reaches an equilibrium frequency of

$$\hat{q} \approx \frac{u}{s[h + (1 - 3h) m_e]}. \tag{A2}$$

As $m_e \to 0$, $\hat{q} \approx u/hs$, as expected from standard theory. Equation A2 shows that somatic mutation increases the equilibrium frequency of inherited mutations when $h > \frac{1}{3}$ and decreases their frequency when $h < \frac{1}{3}$. The effect is, however, usually very small. For a completely dominant deleterious allele, $\hat{q} \approx u/[s(1 - 2m_e)]$. In the limiting case of a completely recessive mutations ($h \to 0$), one can show that

$$\hat{q} \approx \sqrt{\frac{u}{s} + \frac{m_e^2}{4}} - \frac{m_e}{2}. \tag{A3}$$

As $m_e \to 0$, $\hat{q} \approx \sqrt{u/s}$ (HALDANE 1937). As m_e increases, however, q declines, reflecting the fact that selection against inherited recessive mutations is more effective when inherited alleles are "uncovered" by somatic mutation.

Invasion of a diploidy allele: Here I consider the fate of a rare "diploidy" allele (C_2) that appears in a predominantly "haploid" (C_1) population. Before the appearance of C_2, the population is at the mutation-selection equilibrium given by Equation 6. We want to know if C_2 will invade this population.

Because C_2 is very rare, we can linearize the recursions for x_3 and x_4 by ignoring quadratic terms in these variables:

$$\bar{w}x_3' = x_3[(1 - x_2 - u) w_{AA,A\cdot d_{12}} + (1 - r) x_2 w_{Aa,A\cdot d_{12}}] + x_4[(1 - x_2 - u) r w_{Aa,A\cdot d_{12}}]$$

$$\bar{w}x_4' = x_3[r x_2 w_{Aa,a\cdot d_{12}} + u w_{AA,A\cdot d_{12}}] + x_4[(1 - r)(1 - x_2) w_{Aa,a\cdot d_{12}} - u r w_{Aa,A\cdot d_{12}}].$$

If the leading eigenvalue, λ_L, of this system of linear equations is greater than unity, the new diploidy allele invades; if $\lambda_L < 1$, the diploidy allele does not invade. The relevant eigenvalue is

$$\lambda_L = \left(\frac{1}{\bar{w}}\right)\left(K_1 + \frac{K_2}{K_1 - K_3}\right), \tag{A4}$$

where

$$K_1 = (1 - x_2 - u) w_{AA,A\cdot d_{12}} + (1 - r) x_2 w_{Aa,A\cdot d_{12}}$$

$$K_2 = r w_{Aa,A\cdot d_{12}}[r x_2 w_{Aa,a\cdot d_{12}} + u w_{AA,A\cdot d_{12}}]$$

$$K_3 = (1 - r)(1 - x_2) w_{Aa,a\cdot d_{12}} - u r w_{Aa,A\cdot d_{12}},$$

and the w terms are defined as above. When $m_e = 0$, this eigenvalue is equivalent to that given by OTTO and GOLDSTEIN (1992).

Because λ_L is a complicated function of many variables, it is useful to focus on the biologically interesting case of a completely diploid allele ($d_{12} = 1$) that appears in a haploid population ($d_{11} = 0$). In this case, the leading eigenvalue is

$$\lambda_L = 1 + m_e s(1 - 2h) + \frac{u(r - 2hr - h^2 s - hrs + 2h^2 rs)}{r + (1 - r) hs}, \tag{A5}$$

where I again ignore terms of u^2, m_e^2 and um_e and higher.

Examination of this eigenvalue shows that when $m_e = 0$, a diploid allele cannot invade when there is no recombination, as expected (OTTO and GOLDSTEIN 1992). When somatic mutation occurs, however, diploids *can* invade even when $r = 0$. In general, Equation A5 shows that a rare diploidy allele invades when the effective rate of somatic mutation exceeds

$$m_e > \frac{u}{s}\left[\frac{(1 - h) hs}{(1 - 2h)[r + (1 - r) hs]} - 1\right] \tag{A6}$$

and $h < \frac{1}{2}$. Just as with noninterbreeding asexuals (see text, Equation 5), the rate of somatic mutation required to favor diploidy increases with u but decreases with larger s. Most important, Equation A6 shows that when deleterious mutations are partially recessive, there is always some rate of somatic mutation that favors diploidy whether or not the species recombines.