Recombination and the Evolution of Diploidy

Sarah P. Otto and David B. Goldstein

Department of Biological Sciences, Stanford University, Stanford, California 94305

Manuscript received November 8, 1991
Accepted for publication March 19, 1992

ABSTRACT

With two copies of every gene, a diploid organism is able to mask recessive deleterious mutations. In this paper we present the analysis of a two-locus model designed to determine when the masking of deleterious alleles favors the evolution of a dominant diploid phase in organisms that alternate between haploid and diploid phases ("alternation of generations"). It is hypothesized that diploidy will be favored whenever masking occurs ("the masking hypothesis"). Using analytical methods, we confirm that this masking hypothesis is essentially correct under free recombination: as long as the heterozygous expression of deleterious alleles is sufficiently masked by the wild-type allele, diploidy is favored over haploidy. When the rate of recombination is lower, however, diploidy is much less likely to be favored over haploidy. In fact, according to our model, the evolution of diploidy is impossible without significant levels of recombination even when masking is fairly strong.

THE majority of mutations produce inferior or even nonfunctional gene products (LEWONTIN 1974). By having two copies of every gene, diploids are virtually assured of producing at least one normal copy of any protein. As a result, selection should favor diploids over haploids when both groups have the same deleterious allele frequency. This has led many authors to suggest that diploids evolved due to their ability to mask deleterious alleles (CROW and KIMURA 1965; PERROT, RICHERD and VALÉRO 1991; KONDRASHOV and CROW 1991; CHARLESWORTH 1991; MAYNARD SMITH 1978). Diploids carry twice as many alleles as haploids, however, and therefore experience a higher mutation pressure. As a result, the mutational load at equilibrium in a diploid population is twice that in a haploid population, as long as the deleterious allele is not perfectly recessive (CROW and KIMURA 1965).

Considering only individual-level selection and assuming independent assortment, PERROT, RICHERD and VALÉRO (1991) recently showed through computer stimulation that diploids can invade a population whenever the degree of dominance (h) is less than ½, where the magnitude of the reduction below ½ depends on the strength of selection against deleterious alleles. The transition to diploidy, however, may have occurred in relatively "simple" genomes characterized by extremely low rates of recombination. It is therefore worthwhile to consider the evolution of diploidy when the relevant loci are linked. Here we show that, while diploids may invade haploid populations when recombination rates are high (PERROT, RICHERD and VALÉRO 1991), they are never able to do so when recombination rates are very low.

METHODS

We analyze a two-locus model similar to that used in PERROT, VALÉRO, and RICHERD (1991). The viability locus has two alleles (A, a), with A mutating to a at a rate μ. Individuals that carry only mutant alleles (aa or a) have a selective disadvantage of s relative to individuals that carry only the wild-type allele (AA or A). The selective disadvantage of heterozygotes is hs, where h is a measure of dominance. When h is small ($h < ½$), the heterozygote has a viability closer to wild-type individuals than to mutant individuals. In this case, we say that the A allele masks the a allele. The second locus, the ploidy locus, controls the timing of meiosis and consequently the probability that selection occurs during the diploid or the haploid stage. There are two alleles at the ploidy locus. A newly formed zygote C_iC_j undergoes selection as a diploid with probability d_{ij} and as a haploid with probability $(1 - d_{ij})$. We assume that the gametes from individuals having undergone both haploid and diploid selection form a single gamete pool, and unite at random. There are two major differences between our model and that of PERROT, VALÉRO and RICHERD. We include the recombination rate between the loci (r) as a parameter, and allow individuals to undergo selection as diploids with arbitrary probability d_{ij}. These changes generalize the model. Thus, we allow the degree of diploidy to be modified from an arbitrary level to any level and, more importantly, we make no assumptions about the physical arrangement of the viability and ploidy loci. The parameters and life cycle for the two-locus model are summarized in Tables 1 and 2 and in Figure 1.

It is important to note that this formulation assumes the preexistence of a life cycle characterized by an alternation of generations (between haploid and diploid phases) and does not address the evolutionary origin of such a two-stage life cycle. Because it is widely held that the haploid phase was dominant in the first two-stage organisms (RAVEN, EVERT and EICHHORN 1986; MAYNARD SMITH 1978), we cannot use the masking advantage of diploidy to explain the origination of the alternating cycle. Rather, some advantage of a very brief diploid phase is required to explain the origination of the cycle, like the repair of damaged DNA (BERNSTEIN, HOPF and MICHOD 1988). In this paper, we

TABLE 1

The genotypic control of ploidy level

Genotype	Ploidy locus: Proportion diploid[a] (at time of selection)	Ploidy locus: Proportion haploid (at time of selection)
C_1C_1	d_{11}	$(1 - d_{11})$
C_1C_2	d_{12}	$(1 - d_{12})$
C_2C_2	d_{22}	$(1 - d_{22})$

[a] As proportions, $0 \leq d_{11}, d_{12}, d_{22} \leq 1$.

TABLE 2

The genotypic control of selection

Viability locus: Genotype	Viability[a]
AA	1
Aa	$1 - hs$
aa	$1 - s$
A	1
a	$1 - s$

[a] By assumption, selection is acting against the a allele ($0 < s \leq 1$) and $0 \leq h \leq 1$.

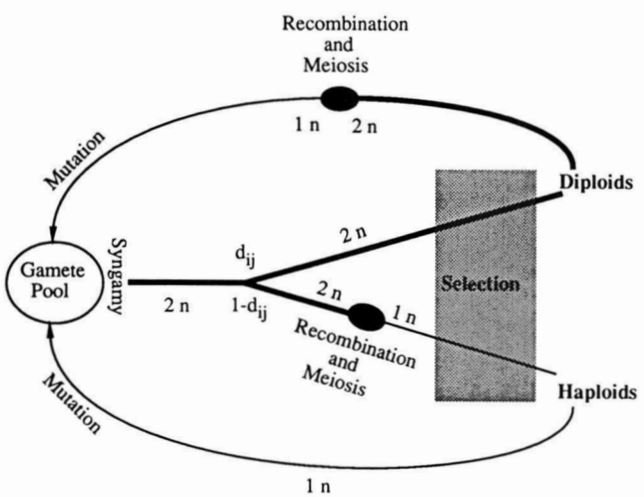

FIGURE 1.—The two-locus model. Gametes carry A or a at the viability locus and either C_1 or C_2 at the ploidy locus. All gamete types unite at random (syngamy). A proportion of these newly formed diploids (d_{ij} for C_iC_j individuals) remain diploid whereas the rest undergo recombination and meiosis becoming haploid. Selection now occurs on all individuals as in Table 2. Following selection, the remaining diploids undergo recombination and meiosis. Finally, all individuals produce gametes, some of which mutate and the cycle begins again. The rate of forward mutation (A to a) is μ. We ignore back mutations (a to A), since the low frequency of the a allele makes these mutations exceedingly rare.

assume the existence of a two-stage life cycle and determine the conditions under which a mutation-selection balance at a viability locus will lead to an increase or decrease of the diploid phase.

RESULTS

Recursion equations which describe the dynamics of the model shown in Figure 1 are given in Appendix 1. From these equations, we have determined that when one ploidy allele is fixed, say C_1, the equilibrium frequency of the deleterious allele at the viability locus is $\mu/[(1 - d_{11})s + d_{11}hs]$, ignoring terms on the order of the square of the mutation rate (μ). (In the limiting case when both $h = 0$ and $d_{11} = 1$, the equilibrium deleterious allele frequency is $\sqrt{\mu/s}$.) We have performed a linear stability analysis (APPENDIX 2) in the vicinity of these equilibria with C_1 fixed to determine the fate of a new ploidy allele (C_2) when it appears in low frequency, say by mutation. If the equilibrium is unstable to the introduction of a new allele that increases the probability of undergoing selection as a diploid ($d_{12} > d_{11}$), then we say that selection favors an increase in the degree of diploidy. This analysis can therefore inform us about the selective forces acting upon ploidy level. Below we present the results of this analysis for general values of the parameter set and for parameters of special interest. In Table 3 we present a summary of the results.

Arbitrary linkage: A new ploidy allele which increases the degree of diploidy ($d_{12} > d_{11}$) will increase if and only if

$$rK_3 > hs[1 - d_{12}(1 - h)] \qquad (1)$$

where

$$K_3 = 1 - 2h + hs - 2d_{12}hs + 2d_{12}h^2s.$$

If condition (1) fails to hold (*e.g.*, if K_3 is negative), only new ploidy alleles that increase the degree of haploidy ($d_{12} < d_{11}$) can increase when rare. This condition is depicted graphically in Figures 2 and 3. Notice first that linkage severely restricts the range of parameters under which diploidy can invade. As the linkage between the loci becomes tighter, the range of the parameter space in the h,s plane which allows the invasion of diploidy shrinks (Figure 2). Another way to consider the result is that as s and h get larger, the recombination rate between the loci required for diploidy to invade, r^*, gets larger (Figure 3).

Free recombination: When $r = \frac{1}{2}$ so that the ploidy and viability loci are unlinked, a new ploidy allele is able to invade if

$$(d_{12} - d_{11})\left(\frac{1}{2} - h - \frac{hs}{2}\right) > 0.$$

Therefore, if the new allele increases the level of diploidy ($d_{12} > d_{11}$), it will invade only if the deleterious mutation is partially recessive (specifically $h < 1/(2 + s)$). Otherwise, with $h > 1/(2 + s)$, an increased level of haploidy is favored by selection. This analytical result is consistent with what PERROT, RICHERD and VALÉRO (1991) found by computer simulation.

Complete linkage: When $r = 0$, *only* alleles that increase the proportion of haploid adults will rise in

TABLE 3

Summary of the analytical results

Case	Leading eigenvalue	Diploidy favored[a]
General	$1 + \frac{\mu(d_{12} - d_{11})K_1}{(1 - d_{11} + hd_{11})K_2}$	$rK_3 > hs(1 - d_{12} + hd_{12})$
$r = 1/2$	$1 - \frac{\mu(d_{12} - d_{11})\,(1 - 2h - hs)}{(1 - d_{11} + hd_{11})\,(1 + s - d_{12}s + d_{12}hs)}$	$h < \frac{1}{(2 + s)}$
$r = 0$	$1 - \frac{h\mu(d_{12} - d_{11})}{(1 - d_{11} + hd_{11})}$	Never[b]
$h = 0$ $(d_{11} < 1)$	$1 + \frac{\mu r(d_{12} - d_{11})}{(1 - d_{11})\,(r + s - d_{12}s - rs + d_{12}rs)}$	Always[b]
$h = 0$ $(d_{11} = 1)$	$1 - \frac{\sqrt{\mu}rs(1 - d_{12})}{(r + s - d_{12}s - rs + d_{12}rs)}$[c]	Always[b]

[a] Diploidy is "favored" whenever the new allele can only invade the population ($\lambda_L > 1$) if it increases the degree of diploidy in the population ($d_{12} > d_{11}$). Haploidy is favored whenever diploidy is not favored. All constants are defined in the text (K_1 and K_2 are defined in APPENDIX 2).

[b] In the limiting case when both $r = 0$ and $h = 0$, selection is globally neutral with respect to ploidy level. Specifically, it can be shown that the fraction (x_1/x_3) does not change over time. Thus, when mutations are rare and deleterious, the ratio of C_1 alleles to C_2 alleles does not change appreciably over time.

[c] This leading eigenvalue was calculated from the recursions linearized near the equilibrium (A2), which explicitly assumes that $(d_{11} = 1)$ and $(h = 0)$.

frequency ($d_{12} < d_{11}$). The percentage of diploid individuals will decrease over time because of selection at the ploidy locus.

Complete masking: When h is exactly zero, masking is complete and alleles that increase the amount of diploidy will always invade. This supports the intuition that when deleterious alleles are *perfectly* recessive diploidy is favored because there is no load cost to diploidy (the haploid and diploid mutational loads are equal when $h = 0$) but a benefit still accrues from masking. This result only holds for recombination rates greater than zero. If both r and h are exactly zero, then there is no selection on the ploidy locus. Most experimental data, however, indicate that there is an appreciable heterozygote effect of deleterious mutations. Data summarized by SIMMONS and CROW (1977) suggest that h lies between 0.01 and 0.03 for recessive lethal mutations ($s \sim 1$) and between 0.3 and 0.5 for mildly deleterious mutations.

Why linkage helps haploids: While diploids mask the deleterious mutations in their genome, such masking enables these mutations to persist over time. Conversely, haploid individuals are more likely to die when they carry a deleterious mutation, removing that mutation from the population, so that haploid populations are more "efficient" at weeding out the mutations that occur (SCUDO 1967). As a result, the equilibrium frequency of mutations ($\mu/[(1 - d_{11})s + d_{11}hs)]$ increases with increasing levels of diploidy (d_{11}). With complete linkage ($r = 0$), haploids only carry their own mutations and only haploids gain from the greater "efficiency" of selection upon them. In essence, haploids are able to carry the haploid load and diploids must carry the diploid load, since recombination is not present to redistribute these loads. Chromosomes carrying the C_1 allele are isolated from chromosomes containing the C_2 allele; they may co-occur within an individual but they will not recombine. Heuristically, then, we can consider the population to be composed of two isolated subpopulations, each carrying a different ploidy allele and a different number of mutant alleles. In such a scenario, arguments based strictly on load can be used to determine whether diploids or haploids will be favored by selection.

Considering a single viability locus, CROW and KIMURA (1965) showed that haploids have a lower mutational load than diploids (recall that load is equal to the reduction in fitness experienced by a group). For an arbitrary degree of diploidy, the mutational load carried by a population composed solely of the C_1 allele is:

$$\mu\,\frac{(1 - d_{11} + 2hd_{11})}{(1 - d_{11} + hd_{11})}$$

which increases as a function of d_{11}, proving that the load is highest (lowest mean fitness) in a completely diploid population. Therefore, when only one locus is considered, haploids always have a lower load and consequently are favored over diploids when the populations are not interbreeding. KONDRASHOV and CROW (1991) have shown, however, that in the case

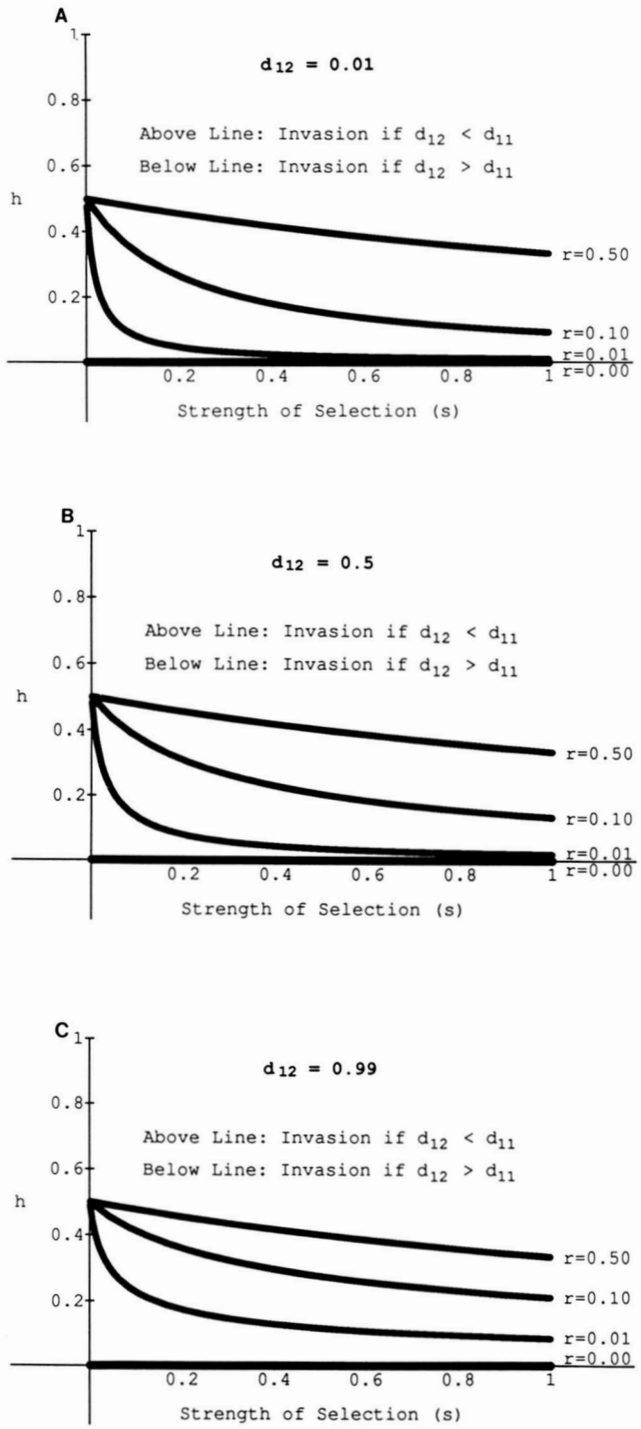

FIGURE 2.—The conditions controlling whether evolution favors a higher or a lower level of diploidy. Condition [1] is depicted graphically for three different new ploidy alleles: $d_{12} = 0.01$ (A), $d_{12} = 0.5$ (B), and $d_{12} = 0.99$ (C). In each graph, we show the line below which only an increase in the level of diploidy is favored (d_{12} must be greater than d_{11} for invasion to occur) and above which only an increase in the level of haploidy is favored (d_{12} must be less than d_{11} for invasion) for four different rates of recombination. Notice that there is a range of parameters (*e.g.*, $r = 0.01$, $s = 0.15$, $h = 0.15$) in which an allele that increases the amount of haploidy (to $d_{12} = 0.01$) can invade but so could an allele that increases the amount of diploidy (to $d_{12} = 0.99$). In such cases, the ploidy level within a population could depend simply on the sequence in which new ploidy alleles appeared in the population.

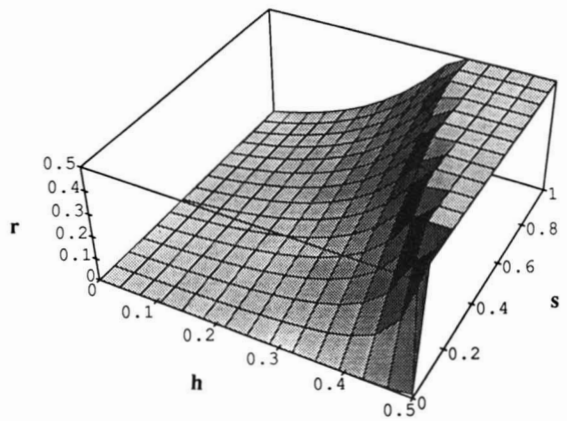

FIGURE 3.—The minimum recombination rate required for up-modification of the diploidy rate. When the new ploidy allele, C_2, is associated with an even probability of producing diploids ($d_{12} = 0.5$) and this corresponds to an up-modification of the ploidy rate ($d_{12} > d_{11}$), then invasion will only occur if $r > r^*$ where r^* is shown. The graph shows only those values of h less than ½. For values of h greater than ½, no value of the recombination rate will allow the evolution of increased diploidy. Note that small selection coefficients (h and s) and large recombination rates favor the evolution of diploidy.

of multiple loci segregating for deleterious alleles, and truncation selection, diploids may have a lower load than haploids. Whenever their analysis applies, and the rate of recombination among all the viability loci and the ploidy locus is zero, then diploidy should be favored.

When r is not strictly zero, the chromosomes are no longer isolated from one another and a comparison of loads no longer suffices to predict the evolutionary fate of a new ploidy allele. Conceptually, recombination transmits more mutations from diploid individuals to haploids and more nonmutant alleles from haploids to diploids, reducing the difference in frequency of mutations between chromosomes carrying different ploidy alleles. We can thus understand the effects of selection and recombination by determining their effects on the associations that develop between the ploidy locus and the viability locus. The arguments described above imply that viability selection drives a positive association between a haploidy allele at the ploidy locus and the nonmutant allele at the selection locus; recombination must be present to break up this association if diploidy is to invade.

We have performed a disequilibrium analysis (UYENOYAMA and BENGTSSON 1989; OTTO 1991) in the vicinity of the equilibrium involving only the C_1 allele to test these heuristics about the role of associations between the ploidy and viability loci. The basic result is that as evolution proceeds the association between the new ploidy allele, C_2, and the non-mutant allele at the viability locus (A) approaches a positive value whenever $d_{12} < d_{11}$ and approaches a negative value whenever $d_{12} > d_{11}$. That is, alleles increasing the probability of undergoing haploid selection always

become positively associated with the superior allele at the viability locus. Furthermore, tight linkage or strong selection either increases the association, or increases the advantage to the haploid allele that results from this association. In fact, even when $r = \frac{1}{2}$, selection creates this association, bringing the critical value of h below $\frac{1}{2}$ and allowing haploidy to be favored even when masking occurs (that is, when $1/(2 + s) < h < \frac{1}{2}$).

Robustness of the results: BENGTSSON (1992) examines a model similar to the one described above. Both his and our study investigate the evolution at a locus controlling ploidy level when this locus is linked to another locus at a mutation-selection balance. However, the two approaches differ in the manner by which ploidy level is modified. BENGTSSON considers a life cycle in which the ploidy locus determines whether haploid individuals are competent to undergo fusion. If fusion does occur, the ensuing organisms experience selection as diploids. If not, the organisms remain haploid throughout the life cycle–never becoming diploid and hence never undergoing recombination. In our paper, all gametes fuse to create diploid zygotes and ploidy level is determined by whether meiosis occurs *before or after* the period of major viability selection. In short, Bengtsson models a modifier of fusion competence and we model a modifier of the timing of meiosis; both modifiers alter the proportion of diploid individuals within a population.

The general conclusion from both models is that evolution of increased diploidy is more likely for loose linkage than for tight linkage, since linkage confers an advantage upon haploids through a positive association between the more haploid allele and the most fit (nonmutant) allele. However, the models differ quantitatively in this respect. In BENGTSSON's model, organisms with a low degree of fusion competence (c) experience a lower rate of recombination, since fewer meiotic events occur (fewer tetrad formations). This lower "effective recombination" creates a higher disequilibrium and hence confers a greater advantage upon haploids in BENGTSSON's model of fusion competence.

Each model applies well to a different biological system. BENGTSSON models a population where the overall degree of recombination among loci is reduced by asexuality, which will occur in haploid Eukaryotes that are only occasionally sexual (such as Rhizopus), and in Prokaryotes which reproduce primarily as haploid asexual organisms that occasionally undergo "sex" (conjugation) during which time the organism is partially diploid. An intriguing result of BENGTSSON's model is that the low levels of recombination induced by asexuality would make it impossible for diploidy to evolve through a masking advantage in Prokaryotes. This provides a possible explanation for why Prokaryotes are haploid, although phylogenetic constraints and a lack of suitable variation might be the appropriate reason. On the other hand, our model applies to all organisms that regularly undergo an alternation of generations (sexual Eukaryotes) in which there is genetic variability for which ploidy phase is dominant. Although we assume that selection only occurs in one phase (either the haploid phase if meiosis is early or the diploid phase if meiosis is late), the qualitative results do not change if selection occurs in both phases (our unpublished data).

The fact that the two models agree completely in their qualitative conclusions provides some evidence for the robustness of the results. The development of associations (disequilibrium) between loci favors the evolution of haploidy and this benefit must be weighed against the masking advantage that favors diploidy in order to determine the fate of new ploidy alleles.

CONCLUSIONS

In summary, under the hypothesis that diploidy evolved in order to mask deleterious recessive alleles, linkage makes the transition from haploidy to diploidy much more difficult. How more than one locus under a mutation-selection balance will affect the evolution of diploidy remains to be seen, but we may speculate that the ratio of tightly linked to loosely linked loci may determine whether diploidy can invade at a ploidy locus. Thus, all else being equal, a larger genome size or a greater division of the genome into chromosomes will favor diploidy, because linkage is, on the whole, looser. For example, in humans, the average level of recombination between any two randomly chosen genes must approximate $\frac{1}{2}$ since there are 23 independently segregating chromosomes. In proto-diploids, however, recombination rates may have been extremely low, perhaps on the order of the mutation rate, as is found in Prokaryotes (MAYNARD SMITH 1988; LEVIN 1988; SELANDER and LEVIN 1980). Such low rates of recombination result whenever there are few chromosomes, few chiasmata per chromosome, or high ratios of apomictic to sexual individuals within a population. If diploidy predated appreciable rates of recombination, then masking deleterious recessive mutations is a much less attractive hypothesis for the evolution of diploidy. Finally, to the extent that masking is important in the evolution of diploidy, there should be a positive correlation between recombination rates and diploidy. While no adequate survey has been made, such a trend has been reported (GHISELIN 1974) and is in agreement with the prevalence of diploidy among higher plants and animals in comparison with protists (BELL 1982).

We thank B. BENGTSSON, A. CLARK, M. FELDMAN, S. TULJAPURKAR, M. UYENOYAMA, and several anonymous reviewers for their

helpful comments on the manuscript. This material is based upon work supported by a National Science Foundation Graduate Fellowship to S.O. and by National Institutes of Health grant GM 28016.

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.

BERNSTEIN, H., F. A. HOPF and R. E. MICHOD, 1988 Is meiotic recombination an adaptation for repairing DNA, producing genetic variation, or both?, pp. 194–211 in *The Evolution of Sex*, edited by R. E. MICHOD and B. R. LEVIN. Sinauer Associates, Sunderland, Mass.

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

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

FELDMAN, M. W., and S. KARLIN, 1971 The evolution of dominance: a direct approach through the theory of linkage and selection. Theor. Popul. Biol. **2:** 482–492.

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

GHISELIN, M. T., 1974 *The Economy of Nature and the Evolution of Sex*. University of California Press, Berkeley.

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

LEVIN, B. R., 1988 The evolution of sex in bacteria, pp. 194–211 in *The Evolution of Sex*, edited by R. E. MICHOD and B. R. LEVIN. Sinauer Associates, Sunderland, Mass.

LEWONTIN, R. C., 1974 *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York.

MAYNARD SMITH, J., 1978 *The Evolution of Sex*. Cambridge University Press, New York.

MAYNARD SMITH, J., 1988 The evolution of recombination, pp. 106–125 in *The Evolution of Sex*, edited by R. E. MICHOD and B. R. LEVIN. Sinauer Associates, Sunderland, Mass.

OTTO, S. P., 1991 On evolution under sexual and viability selection: a two-locus diploid model. Evolution **45:** 1443–1457.

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

RAVEN, P. H., R. F. EVERT and S. E. EICHHORN, 1986 *Biology of Plants*. Worth Publishers, New York.

SCUDO, F. M., 1967 Selection on both haplo and diplophase. Genetics **56:** 693–704.

SELANDER, R. K., and B. R. LEVIN, 1980 Genetic diversity and structure in populations of *Escherichia coli*. Science **210:** 545–547.

SIMMONS, M. J., and J. F. CROW, 1977 Mutations affecting fitness in Drosophila populations. Annu. Rev. Genet. **11:** 49–78.

UYENOYAMA, M. K., and B. O. BENGTSSON, 1989 On the origin of meiotic reproduction: a genetic modifier model. Genetics **123:** 873–885.

Communicating editor: A. G. CLARK

APPENDIX 1

We census at the gamete stage and define:

x_1 = frequency (AC_1) $\quad x_2$ = frequency (aC_1)

x_3 = frequency (AC_2) $\quad x_4$ = frequency (aC_2),

where $x_1 + x_2 + x_3 + x_4 = 1$. The general recursions are:

$$\begin{aligned}
Tx_{1'} &= (1-\mu)\{x_1^2 + x_1x_2(1 - hsd_{11}) + x_1x_3 \\
&\quad + (1-r)x_1x_4(1 - hsd_{12}) + r\,x_2x_3(1 - hsd_{12})\} \\
Tx_{2'} &= \mu\{x_1^2 + x_1x_2(1 - hsd_{11}) + x_1x_3 \\
&\quad + (1-r)x_1x_4(1 - hsd_{12}) + rx_2x_3(1 - hsd_{12})\} \\
&\quad + \{x_1x_2(1 - s + sd_{11} - hsd_{11}) + rx_1x_4(1 - s \\
&\quad + sd_{12} - hsd_{12}) + x_2^2(1 - s) + (1-r)x_2x_3 \\
&\quad \cdot (1 - s + sd_{12} - hsd_{12}) + x_2x_4(1 - s)\} \\
Tx_{3'} &= (1-\mu)\{x_1x_3 + rx_1x_4(1 - hsd_{12}) + x_2x_3(1-r) \\
&\quad \cdot (1 - hsd_{12}) + x_3^2 + x_3x_4(1 - hsd_{22})\} \\
Tx_{4'} &= \mu\{x_1x_3 + rx_1x_4(1 - hsd_{12}) + x_2x_3(1-r) \\
&\quad \cdot (1 - hsd_{12}) + x_3^2 + x_3x_4(1 - hsd_{22})\} \\
&\quad + \{x_1x_4(1-r)(1 - s + sd_{12} - hsd_{12}) \\
&\quad + x_2x_3(1 - s + d_{12}s - hsd_{12})r + x_2x_4(1 - s) \\
&\quad + x_3x_4(1 - s + sd_{22} - hsd_{22}) + x_4^2(1 - s)\}
\end{aligned}$$

where T is the sum of the right hand sides and is equal to the mean fitness. We assume, as is standard, that the deleterious allele frequency is sufficiently small that back mutation may be ignored. Interestingly, these recursions are clearly analogous to those describing the evolution of dominance (FELDMAN and KARLIN 1971). With the C_1 allele fixed ($x_3 = x_4 = 0$), we have that, at equilibrium:

$$\begin{aligned}
x_1 &= 1 - x_2 \\
x_2 &= \frac{\mu}{s(1 - d_{11}) + hsd_{11}}.
\end{aligned} \tag{A1}$$

This equilibrium is invalid if the mutation (a) is completely recessive ($h = 0$) and all individuals are diploid at the time of selection ($d_{11} = 1$). In this case, the equilibrium configuration of the population is:

$$\begin{aligned}
x_1 &= 1 - x_2 \\
x_2 &= \sqrt{\frac{\mu}{s}}
\end{aligned} \tag{A2}$$

as first noted by HALDANE (1937). Since this equilibrium is a limiting case (both $h = 0$ and $d_{11} = 1$ must hold), we will concentrate in APPENDIX 2 on the stability of the equilibrium described by (A1).

APPENDIX 2

At the equilibrium with C_1 fixed ([A1]), we consider the introduction of a small number of individuals carrying a new allele, C_2. Since the perturbation is small (x_3 and x_4 are small), we ignore any quadratic

terms in these variables within the recursions. The linearized recursions so produced are:

$$x_3' = (1 - \mu) \cdot \frac{x_3 - hsd_{12}x_3\mu/[s(1 - d_{11}(1 - h))] + r(1 - hsd_{12})D}{T}$$

$$x_4' = \mu \frac{x_3 - hsd_{12}x_3\mu/[s(1 - d_{11}(1 - h))] + r(1 - hsd_{12})D}{T} + \frac{x_4(1 - s) + x_4sd_{12}(1 - h)(1 - \mu/[s(1 - d_{11}(1 - h))]) - r(1 - s - hsd_{12} + sd_{12})D}{T}$$

where

$$D = x_1x_4 - x_2x_3 = \left(1 - \frac{\mu}{s(1 - d_{11}(1 - h))}\right)x_4 - \frac{\mu}{s(1 - d_{11}(1 - h))}x_3.$$

We find that the leading eigenvalue (λ_L) of this linearized system equals:

$$\lambda_L = 1 + \frac{\mu(d_{12} - d_{11})K_1}{(1 - d_{11} + hd_{11})K_2}$$

where

$$K_1 = r - 2hr - hs + hd_{12}s - h^2d_{12}s + hrs - 2hd_{12}rs + 2h^2d_{12}rs$$

$$K_2 = r + s - d_{12}s + hd_{12}s - rs + d_{12}rs - hd_{12}rs.$$

When the leading eigenvalue is greater than unity in magnitude, the new allele, C_2, increases in frequency and invades the original population. Conversely, whenever the leading eigenvalue is less than one in absolute value, the new allele decreases in frequency over time. Upon further examination of λ_L, it can be seen that when K_1 is positive, only new alleles that increase the amount of diploidy ($d_{12} > d_{11}$) will invade, while when K_1 is negative, only new alleles that reduce the amount of diploidy ($d_{12} < d_{11}$) will invade the population. Hence, selection favors increased diploidy only when $K_1 > 0$, which requires that:

$$rK_3 > hs(1 - d_{12} + hd_{12})$$

where

$$K_3 = 1 - 2h + hs - 2d_{12}hs + 2d_{12}h^2s.$$

Clearly K_3 must be positive for diploidy to increase. Invasion criteria for parameter sets of special interest are given in Table 3.