An Adaptive Hypothesis for the Evolution of the *Y* Chromosome

H. Allen Orr and Yuseob Kim

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

Manuscript received May 28, 1998
Accepted for publication September 16, 1998

ABSTRACT

Population geneticists remain unsure of the forces driving the evolution of *Y* chromosomes. Here we consider the possibility that the degeneration of the *Y* reflects its inability to evolve adaptively. Because the overwhelming majority of favorable mutations on a nonrecombining proto-*Y* suffer a zero probability of fixation, the fitness of the *Y* must lag far behind that of the recombining *X*. At some point, this disparity will grow so large that selection favors an increase in the expression of (fit) *X*-linked alleles and a decrease in the expression of (unfit) *Y*-linked alleles. Our calculations suggest that this process acts far more rapidly than hitchhiking-induced erosion of the *Y* and at least as rapidly as the fixation of deleterious alleles on the *Y* by background selection. Most important, this hypothesis can explain the evolution of *Y* chromosomes in taxa such as Drosophila that have very large population sizes.

GIVEN the ubiquity of *Y* chromosomes (Bull 1983), it is surprising that we remain uncertain of the forces driving the evolution of these largely inert chromosomes. There have, of course, been many attempts to explain *Y* evolution. Virtually all assume that the *Y* was, ancestrally, a normal recombining homolog that harbored functional genes, an assumption for which there is good evidence (*e.g.*, Goodfellow *et al.* 1983). Moreover, all of these explanations depend on the fact that at some point the proto-*Y* stopped recombining with its proto-*X* homolog.

Muller (1918) offered the first theory of *Y* degeneration. He noted that, because the *Y* is nonrecombining, *Y*-linked mutations remain permanently heterozygous. Recessive mutations are therefore unopposed by selection and so might slowly accumulate on the *Y*, ultimately rendering the chromosome nonfunctional. Unfortunately, Fisher (1935) showed that Muller's intuition was mistaken: the inevitable unmasking of *Y*-linked mutations by homologous *X*-linked mutations suffices to keep *Y*-linked deleterious alleles at low frequencies.

Although more modern theories of *Y* evolution have fared somewhat better, none is wholly satisfactory. Nei (1970), for instance, showed that Muller's argument can be rescued in sufficiently small populations because deleterious *Y*-linked alleles can drift to fixation. But as Charlesworth (1978) pointed out, this theory, while formally sound, suffers two problems: (1) Because the relevant effective size is that of the *species*, it seems unlikely that the theory can apply to many taxa. Indeed, Nei's theory requires that the effective size be less than ~10,000, a figure several orders of magnitude smaller than that estimated from molecular studies of Drosophila (Moriyama and Powell 1996), a group that harbors degenerate *Y* chromosomes. (2) Dosage compensation suggests that the *Y* did *not* degenerate due to the accumulation of completely recessive alleles: there can be no selection for dosage compensation if individuals carrying a single functional copy of a gene enjoy the same fitness as those carrying two functional copies.

Charlesworth (1978) instead proposed that *Y* degeneration results from Muller's ratchet, the irreversible accumulation of deleterious alleles due to stochastic loss of mutation-free nonrecombining chromosomes. As then understood, this theory differed profoundly from Nei's: Muller's ratchet caused only an increase in the mean number of deleterious alleles per *Y*, not fixation of *Y*-linked mutations. Recent work, however, has shown that this interpretation is incorrect, at least for haploid chromosomes like the *Y*. In such cases, each turn of Muller's ratchet *is* associated with fixation of a deleterious mutation (Charlesworth and Charlesworth 1997). Muller's ratchet can then, in principle, cause inactivation of *Y*-linked genes. Unfortunately, however, Muller's ratchet turns at an exceedingly slow rate in large species (Stephan *et al.* 1993). Using parameter values from Drosophila, Charlesworth (1996) estimates that, in a species of population 1 million, 10^{30} generations pass between *each* turn of the ratchet. Even in fast-breeding organisms (*e.g.*, one generation per week), there therefore has not been enough time since the origin of the universe for a single turn of the ratchet (but see Gessler 1995). Thus, while Charlesworth (1996) and Charlesworth and Charlesworth (1997) conclude that Muller's ratchet might play some role in *Y* evolution in taxa such as mammals, it would not appear an effective force in large species like Drosophila.

Corresponding author: H. Allen Orr, Department of Biology, University of Rochester, Rochester, NY 14627.
E-mail: haorr@urhep.pas.rochester.edu

Our problem therefore is simple: given that *Y* chromosomes are as common in Drosophila as in mammals, we require a theory of *Y* evolution that holds in small *and* large taxa.

Rice proposed such a theory in 1987. Under his hypothesis, the *Y* degenerates as a side effect of the substitution of favorable mutations. Because most *Y* chromosomes at mutation-selection balance carry one or more deleterious alleles, favorable mutations often arise on mutation-bearing *Y*'s. When favorable mutations sweep to fixation, they may drag deleterious alleles at other *Y*-linked loci to fixation, gradually eroding *Y* function (Rice 1987). Unfortunately, this process is also excruciatingly slow. The reason is that the vast majority of favorable mutations that go to fixation arise on mutation-free chromosomes and so cause no hitchhiking (Manning and Thompson 1984; Peck 1994).

Charlesworth (1996) has recently proposed yet another hypothesis for the degeneration of the *Y*. He argues that background selection—the reduction in effective population size caused by deleterious mutations on nonrecombining chromosomes—allows slightly deleterious alleles to drift to fixation. In a recent review, Charlesworth (1996) concludes that background selection represents the most plausible explanation of *Y* degeneration, at least in large taxa where Muller's ratchet is ineffective. While it seems clear that background selection must act in regions of no recombination, and thus that background selection must contribute to the degeneration of the *Y*, this theory suffers the same problem as the above: it is slow, especially in large species (Charlesworth 1996).

Here we consider a hypothesis for *Y* evolution that, we believe, escapes many of the problems besetting the above theories. Following Rice (1996), we argue that the degeneration of the *Y* may be a consequence of its inability to evolve adaptively. Because, as noted above, the overwhelming majority of favorable mutations appear on *Y* chromosomes bearing one or more deleterious mutations (Manning and Thompson 1984), virtually all *Y*-linked favorable mutations suffer a zero probability of fixation, a situation Peck (1994) deemed the "ruby-in-the-rubbish" problem (see also Fisher 1930, p. 122). Nonrecombining chromosomes thus suffer an extremely slow rate of adaptive evolution compared to recombining ones. Because the fitness of the *Y* must lag far behind that of the *X*, there will be selection over time to increase the expression of fit *X*-linked alleles and to decrease the expression of unfit *Y*-linked alleles.

Although this idea was previously sketched by Rice (1996), he did not present calculations on the efficiency of adaptation-driven degeneration of the *Y*. Here we calculate substitution rates for favorable mutations on the proto-*Y* *vs.* proto-*X* and, more importantly, contrast the efficiency of the ruby-in-the-rubbish process for *Y* evolution with that due to hitchhiking and fixation of deleterious alleles by background selection. Our results suggest that the ruby-in-the-rubbish process is a potent force for *Y* evolution. Most importantly, this force can drive *Y* evolution in species of arbitrarily large size. Indeed, it is *more* effective in larger taxa.

THE MODEL

Adaptive evolution on the *X vs. Y*: We consider adaptation in a diploid population of size *N*. We assume that evolution is due to the substitution of favorable new mutations that arise at a per-chromosome rate of U_f. These alleles have a distribution, $f(s_f)$, of favorable heterozygous effects, with the mean mutation enjoying an advantage of $\bar{s}_f$. On a recombining autosome, the average favorable mutation enjoys a probability of fixation of $\sim 2\bar{s}_f$ (with reasonably weak selection), and the rate of adaptive substitution is therefore $k_A \approx (2NU_f)(2\bar{s}_f) = 4NU_f\bar{s}_f$.

On an equivalent-sized proto-*X* chromosome, the adaptive substitution rate is

$$k_X \approx (\tfrac{3}{2}NU_f)(2\bar{s}_f) = 3NU_f\bar{s}_f. \quad (1)$$

Because the *Y* is not yet degenerate, we assume that *X*-linked alleles are selected only for their heterozygous, not hemizygous, effects. Equation 1 thus differs from previous formulae for adaptive substitution rates on the *X* (*e.g.*, Charlesworth *et al.* 1987).

On a nonrecombining proto-*Y*, probabilities of fixation are reduced because favorable mutations must often fall onto chromosomes bearing deleterious mutations. We can write

$$k_Y \approx (\tfrac{1}{2}NU_f)(2s_{eff}) = NU_f s_{eff}, \quad (2)$$

where $2s_{eff}$ is the average probability of fixation enjoyed by favorable mutations appearing on the proto-*Y*.

To find s_{eff}, first consider the case in which all deleterious mutations have the same effect, s_d. At mutation-selection balance, the number of deleterious alleles per chromosome is Poisson distributed: $P_k = e^{-U_d/s_d}(U_d/s_d)^k/k!$, where U_d is the per-*Y*-chromosome deleterious mutation rate (Kimura and Maruyama 1966). If $\bar{s}_{fk}$ is the mean net selection coefficient enjoyed by favorable mutations arising on chromosomes carrying *k* deleterious alleles, then the average favorable mutation experiences an initial advantage of about

$$\begin{aligned} &\approx P_0\bar{s}_{f0} + P_1\bar{s}_{f1} + P_2\bar{s}_{f2} + \ldots \\ &\approx \sum_{k=0}^{\infty} P_k \int_{ks_d}^{\infty} (s_f - ks_d)\, f(s_f)\, ds_f, \end{aligned} \quad (3)$$

where we assume multiplicative fitness effects among deleterious mutations [and that $1 - (1 - s_d)^k \approx ks_d$] and ignore those cases where a mutation's favorable effect is less than the chromosome's net deleterious effects, because such chromosomes cannot contribute to distant future generations.

When $\bar{s}_f \ll s_d$, adaptive evolution is constrained by the size of the zero class, P_0, and Equation 3 reduces to

$$s_{eff} \approx e^{-U_d/s_d}\,\bar{s}_f. \quad (4)$$

Peck (1994) and Barton (1995, p. 830) have previously shown that, as expected, twice Equation 4 gives the correct probability of fixation for favorable mutations arising on nonrecombining chromosomes [see especially Peck (1994), Table 1, which includes the results of computer simulations]. The intuitive reason is that, when $\bar{s}_f \ll s_d$, favorable mutations arising on mutation-free chromosomes enjoy a "normal" probability of fixation of $2\bar{s}_f$, while those arising on deleterious mutation-bearing chromosomes suffer a zero probability of fixation; the net probability of fixation is therefore $2\bar{s}_f$, weighted by the frequency of mutation-free chromosomes ($\exp(-U_d/s_d)$).

The biologically important point is that the rate of adaptive substitution on the *X* relative to the *Y* is

$$\frac{k_X}{k_Y} = \frac{3\bar{s}_f}{s_{eff}} \approx 3e^{U_d/s_d}. \quad (5)$$

For plausible parameter values, this ratio is very large. If, for instance, $U_d = 0.1$, $s_d = 0.02$ (Peck 1994; Charlesworth 1996), and $\bar{s}_f = 0.001$, we get $k_X/k_Y = 445$. With weak selection, this ratio has the convenient property of being nearly independent of $\bar{s}_f$ (Equation 5), a poorly known quantity.

So far we have only allowed for a distribution of favorable effects. If we also allow for a distribution of deleterious effects, Equations 4 and 5 remain reasonably accurate if we replace s_d with the harmonic mean $\bar{s}_H$ of deleterious effects among newly arising mutations [see appendix and Charlesworth (1996)]. [This approximation is good as long as there are not too many mutations of very small effect (Charlesworth 1996). We have confirmed the accuracy of this approximation in computer simulations (results not shown).] Crow and Simmons's (1983) review of the Drosophila data suggests that $\bar{s}_H = 0.02$ (see also Charlesworth 1996). Thus, with $U_d = 0.1$, we still obtain $k_X/k_Y \approx 445$, and the *Y* chromosome lags far behind the rapidly evolving *X*.

In fact, it is easy to calculate just how much the *Y* lags behind. If $w_{X/Y}$ gives the ratio of *X* to *Y* fitness and fitness is multiplicative across loci, the rate at which this fitness difference grows is

$$\frac{d \ln w_{X/Y}}{dt} = 3NU_f s_f^2 - NU_f s_{eff} s_f \approx 3NU_f \bar{s}_f^2, \quad (6)$$

where the approximation assumes that adaptation on the *Y* is negligibly slow.

Comparison with hitchhiking and background selection: The ruby-in-the-rubbish process—and hitchhiking and the fixation of deleterious alleles by background selection—will surely contribute to a widening gap in fitness between the *X* and *Y*. The important question is quantitative: Is this effect smaller or larger or on par with these other forces? Here we compare these processes.

First consider hitchhiking. Because hitchhiking's effect is proportional to the rate of adaptive substitution on the *Y* ($\approx NU_f s_{eff}$), and to the average number, $\bar{k}$, of deleterious mutations that get dragged to fixation with each substitution (including those cases in which none is dragged along), the fitness of the *Y* relative to *X* declines at a rate of

$$\frac{d \ln w_{W/Y}}{dt} \approx NU_f s_{eff} \bar{k} \bar{s}_H. \quad (7)$$

This rough calculation considers only those loci at which deleterious mutations get fixed; *i.e.*, we generously ignore the fact that favorable substitutions improve the fitness of some loci and of the *Y* chromosome as a whole. In Equation 7 $\bar{k}$ will be close to

$$\bar{k} \approx \frac{\sum_{k=0}^{\infty} k\,P_k \int_{k\bar{s}_H}^{\infty} (s_f - k\bar{s}_H) f(s_f)\,ds_f}{\sum_{k=0}^{\infty} P_k \int_{k\bar{s}_H}^{\infty} (s_f - k\bar{s}_H) f(s_f)\,ds_f}, \quad (8)$$

i.e., the number of deleterious mutations per chromosome class ($k = 0, 1, \ldots$) weighted by the mean fitness advantage of such a chromosome (conditional on a net advantage).

Substituting Charlesworth's parameter values into Equations 7 and 8 and allowing favorable effects to be exponential with $\bar{s}_f = 0.001$, we find that the ruby-in-the-rubbish process causes a 10^9-fold faster divergence in *X vs. Y* fitness than hitchhiking. (It is worth noting that this comparison is independent of the rate of mutation to favorable alleles as both processes depend equally on U_f.) Although this numerology should not be taken too literally, the effect of hitchhiking is clearly small relative to the ruby-in-the-rubbish. The reason is simple: when $\bar{s}_f < \bar{s}_H$, $\bar{k}$ must be quite small because almost all favorable substitutions involve *Y* chromosomes that are free of deleterious mutations; *i.e.*, adaptive evolution is almost entirely constrained to the $k = 0$ class (Peck 1994). Thus, as Charlesworth (1996, p. 155) concluded, hitchhiking is "almost certainly less important than was originally envisaged."

Now consider the fixation of deleterious alleles by background selection. By reducing the effective size of nonrecombining chromosomes, background selection against strongly deleterious alleles can allow very slightly deleterious mutations to drift to fixation. Such alleles must have extremely small deleterious effects or fixation is essentially impossible. Thus, as Charlesworth (1994, 1996) emphasized, these alleles form a distinct class from the more strongly deleterious ones actually causing background selection. (Charlesworth suggests that the former's effects are at least two orders of magnitude smaller than the latter's.) To emphasize this distinction, we write s_s for the strength of selection against,

and U_s for the per-chromosome rate of mutation to, slightly deleterious alleles.

Fixation of deleterious alleles by background selection causes the fitness of the *Y* relative to the *X* to decline at a rate of roughly

$$\frac{d \ln w_{X/Y}}{dt} \approx \frac{P_0 N U_s s_s^2}{\exp(P_0 N s_s) - 1}. \quad (9)$$

This result is essentially equivalent to Equation 3 in Charlesworth (1996), except that we express time in units of generations and assume an even sex ratio ($N = 2N_m$). From Equations 6 and 9, the ruby-in-the-rubbish therefore will cause faster divergence in *X vs. Y* fitness than background selection when

$$3U_f \bar{s}_f^2 > \frac{P_0 \, s_s^2 U_s}{\exp(P_0 N s_s) - 1}. \quad (10)$$

Substituting Charlesworth's favored parameter values, the ruby-in-the-rubbish effect exceeds that due to fixation of deleterious alleles when $\bar{s}_f^2 \, U_f / U_s > 10^{-11}$. Thus, if $\bar{s}_f = 0.001$ and favorable mutations are 1000-fold rarer than those to very slightly deleterious alleles, the ruby-in-the-rubbish process is two orders of magnitude faster than fixation by background selection. This could well be an underestimate for two reasons. First, Charlesworth's (1996) parameter values were explicitly chosen to find the *maximum* rate of decline in *Y* fitness due to fixation of deleterious alleles. In particular, fixation of deleterious alleles must occur at far lower rates in populations that are larger than those assumed by Charlesworth. Second, because the number of mutations that can occur at any locus is limited [*i.e.*, Gillespie's (1991) granularity argument], few mutations may fall into the required very slightly deleterious class at any locus.

DISCUSSION

The present hypothesis may seem counterintuitive for two reasons. First, because the ruby-in-the-rubbish process does not cause fixation of nonfunctional alleles on the *Y*, it may seem to fail to explain that feature of the *Y* that most needs explaining—its genetic inertness. It is important to understand, however, that the present theory posits a two-step process: (1) the fitness of the proto-*Y* lags behind that of the proto-*X* and (2) at some point, this disparity grows so large that it pays to increase the expression of *X*-linked genes at the expense of *Y*-linked genes in males (this will be true as long as the favorable mutations substituted on the *X* are not completely dominant). This requirement of subsequent direct selection for *Y*-inactivation leaves the present hypothesis in the same situation as Charlesworth's (1978) original Muller's ratchet theory. As then understood, Muller's ratchet did not involve fixation of deleterious mutations, only an increase in the mean number of mutations per *Y*. Subsequent selection was thus invoked to explain the heightened expression of the *X* chromosome and the "shutting off" of the *Y* chromosome (Charlesworth 1978).

Similarly, the current leading theory of *Y* evolution—fixation of deleterious alleles by background selection—also likely requires subsequent direct selection to explain *Y*-inactivation. As Charlesworth (1996) emphasizes, background selection can only cause fixation of mutations of very slightly deleterious effect. *Y*-inactivation, therefore could evolve in two ways. First, multiple mutations might get fixed in each *Y*-linked gene, ultimately inactivating it. This gene-by-gene evolution of inactivation, and of dosage compensation, would be exceedingly slow. If on average 3 very slightly deleterious substitutions were required to inactivate a functional gene, 3000–6000 background-selection-induced deleterious substitutions would be required to produce a modern, inactive *Drosophila melanogaster Y* (where we assume that each of Bridges's bands corresponds to one to two loci). Alternatively, fixation of deleterious alleles by background selection might cause the evolution of *X* dosage compensation and of *Y*-inactivation on a chromosome "block by block" basis; the fixation of slightly deleterious alleles then "creates a selection pressure for increasing the activity of *X*-linked loci at the expense of the homologous *Y*-linked loci, leading eventually to the evolution of inactive *Y*-linked loci" (Charlesworth 1996, p. 155). This is the same two-step scenario invoked by the present theory.

The intuition that rapid adaptive evolution of the *X* cannot explain the degeneration of the *Y* is perhaps best dispelled by the following exercise. Imagine that the *X* and the *Y* accumulate favorable substitutions at the same rapid rate. Now imagine that, after a long period of time, all those *Y*-linked genes at which favorable substitutions occurred suddenly revert to their ancestral alleles. These now-unfit revertants obviously behave as deleterious mutations, and we would expect all the evolutionary consequences of an accumulation of such deleterious alleles to follow, *e.g.*, selection for *Y*-inactivation. But clearly these *Y*-linked alleles are just as deleterious whether they just mutated or were there all along.

Second, it may seem that the ruby-in-the-rubbish process, depending as it does on adaptive evolution, would be too slow to account for *Y* evolution. There are several points to bear in mind. First, because selection to inactivate the *Y* reflects the fact that *Y* chromosome fitness lags too far behind the *X*, the relevant substitution rates are those for the *entire Y vs.* the entire *X* (at least if inactivation and dosage compensation evolve block by block). Second, it is important to note that a popular alternative explanation of *Y* evolution—Rice's hitchhiking model—*also* depends on adaptive evolution. Indeed, as emphasized above, the ruby-in-the-rubbish process likely acts far faster than hitchhiking because the former

depends on that great majority of favorable mutations that are fixed on the *X* but not on the *Y*, while the latter is limited to those rare favorable mutations that are fixed on the *Y* (Equations 6 and 7). Last, the leading theory of *Y* evolution—background selection—also depends on a slow process, the fixation of deleterious mutations. Although perhaps occurring at an appreciable rate in small species, fixation of deleterious alleles remains very difficult in large species. But even if deleterious substitutions on the *Y do* occur at an appreciable pace—for example, even faster than adaptive substitution on the *X*—the total *fitness* effects of these *Y*-linked substitutions may be small relative to adaptive ones on the *X*. Background selection, after all, is constrained to fixing alleles of very slight (deleterious) effect, while adaptive evolution on the *X* might fix alleles of quite large (favorable) effect. (Technically, $d \ln w_{X/Y}/dt = (d \ln w_{X/Y}/dk)\,(dk/dt)$, and any comparison of the efficacy of two forces must consider both terms, not merely the second, the substitution rate.) Indeed, our calculations suggest that, for plausible parameter values, the ruby-in-the-rubbish process causes faster divergence in the fitness of the *X* and the *Y* than does background selection.

But too much should not be made of such quantitative comparisons. The rates at which ruby-in-the-rubbish *vs.* fixation of deleterious alleles by background selection act depend on several poorly known parameters, especially the rate of mutation to favorable alleles, population size, and the mean advantage enjoyed by favorable mutations. In reality, *both* processes must almost certainly contribute to the evolution of *Y* chromosomes from functional homologs. These processes will act at a combined rate of about

$$\frac{d \ln w_{X/Y}}{dt} \approx N\left[3U_f\bar{s}_f^2 + \frac{P_0U_ss_s^2}{\exp(P_0Ns_s) - 1}\right]. \qquad (11)$$

(One can also easily imagine an interaction between the two processes wherein substitution of a deleterious allele on the *Y* drives adaptive substitution on the *X* to cover these deleterious effects, and so on.) Indeed, it is important to note that both processes are simple consequences of background selection. That is, both the ruby-in-the-rubbish and fixation of slightly deleterious alleles reflect the fact that, at mutation-selection balance, deleterious mutations reduce the effective population sizes of nonrecombining chromosomes.

Despite this shared basis, however, the two processes show quite different dynamics. While the rate of fixation of deleterious alleles by background selection slows with increasing population size, the ruby-in-the-rubbish effect actually accelerates (Equation 6). The reason for this difference is simple. As population size increases, the size of the zero-class, P_0N, increases and so the upward drift of slightly deleterious mutations within this class grows more difficult. Consequently, the rate of fixation of deleterious mutations declines almost linearly with *N* (Equation 9). But while the rate of adaptive evolution on the *Y* is more or less constrained by the size of the zero class, the rate of adaptive evolution on the *X* is *not*. Consequently, the rate at which *X* and *Y* fitness diverge grows even faster as *N* increases (Equation 6).

The biologically important point is obvious: the ruby-in-the-rubbish process can cause the evolution of *Y* chromosomes in species of arbitrarily large size.

We thank Brian Charlesworth, Andy Clark, Corbin Jones, Daven Presgraves, Wolfgang Stephan, and two anonymous reviewers for very helpful comments. This work was supported by the National Institutes of Health grant GM-51932 and by the David and Lucile Packard Foundation.

LITERATURE CITED

Barton, N. H., 1995 Linkage and the limits of natural selection. Genetics **140:** 821–841.

Bull, J. J., 1983 *Evolution of Sex Determining Mechanisms*. Benjamin-Cummings, Menlo Park, CA.

Charlesworth, B., 1978 Model of evolution of Y chromosomes and dosage compensation. Proc. Natl. Acad. Sci. USA **75:** 5618–5622.

Charlesworth, B., 1994 The effect of background selection against deleterious mutations on weakly selected, linked variants. Genet. Res. **63:** 213–227.

Charlesworth, B., 1996 The evolution of chromosomal sex determination and dosage compensation. Curr. Biol. **6:** 149–162.

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

Charlesworth, B., J. A. Coyne and N. Barton, 1987 The relative rates of evolution of sex chromosomes and autosomes. Amer. Natur. **130:** 113–146.

Crow, J. F., and M. J. Simmons, 1983 The mutation load in *Drosophila*, pp. 1–35 in *The Biology and Genetics of Drosophila*. Academic Press, London.

Fisher, R. A., 1930 *The Genetical Theory of Natural Selection*. Oxford University Press, Oxford.

Fisher, R. A., 1935 The sheltering of lethals. Biol. Rev. **6:** 446–455.

Gessler, D. D. G., 1995 The constraints of finite size in asexual populations and the rate of the ratchet. Genet. Res. **66:** 241–253.

Gillespie, J. H., 1991 *The Causes of Molecular Evolution*. Oxford University Press, Oxford.

Goodfellow, P., G. Banting, D. Sheer, H. Ropers, A. Caine *et al.*, 1983 Genetic evidence that a Y-linked gene in man is homologous to a gene on the X chromosome. Nature **302:** 346–349.

Kimura, M., and T. Maruyama, 1966 The mutational load with epistatic interactions in fitness. Genetics **54:** 1303–1312.

Manning, J. T., and D. J. Thompson, 1984 Muller's ratchet and the accumulation of favourable mutations. Acta Biotheor. **33:** 219–225.

Moriyama, E. N., and J. R. Powell, 1996 Intraspecific nuclear DNA variation in Drosophila. Mol. Biol. Evol. **13:** 261–277.

Muller, H. J., 1918 Genetic variability, twin hybrids and constant hybrids, in a case of balanced lethal factors. Genetics **3:** 422–499.

Nei, M., 1970 Accumulation of nonfunctional genes on sheltered chromosomes. Am. Nat. **104:** 311–322.

Peck, J., 1994 A ruby in the rubbish: beneficial mutations, deleterious mutations, and the evolution of sex. Genetics **137:** 597–606.

Rice, W. R., 1987 Genetic hitch-hiking and the evolution of reduced genetic activity of the Y sex chromosome. Genetics **116:** 161–167.

Rice, W. R., 1996 Evolution of the Y sex chromosome in animals. Bioscience **46:** 331–343.

Stephan, W., L. Chao and J. G. Smale, 1993 The advance of Muller's ratchet in a haploid asexual population: approximate solutions based on diffusion theory. Genet. Res. **61:** 225–232.

Communicating editor: A. G. Clark

APPENDIX: THE ROLE OF THE HARMONIC MEAN, $\bar{s}_H$

When there is a distribution of deleterious effects among new mutations, it is easy to show that the mean fitness effect at mutation-selection balance approximately equals the harmonic mean of deleterious effects among new mutants. Let $f(s_d)$ be the probability density of deleterious heterozygous effects among newly arising mutations. At mutation-selection equilibrium, the fitness effect of the average deleterious mutation is

$$E[\hat{s}_d] = \frac{\int_0^1 s_d \hat{q} f(s_d) ds_d}{\int_0^1 \hat{q} f(s_d) ds_d}, \quad \text{(A1)}$$

where $\hat{q}$ gives the frequency of a deleterious mutation at mutation-selection balance. But at equilibrium, $\hat{q} \approx u/s_d$, where the approximation assumes that $s_d \gg u$. Substituting in A1 and simplifying, we get

$$E[\hat{s}_d] = \frac{1}{\int_0^1 (1/s_d) f(s_d) ds_d}, \quad \text{(A2)}$$

which is the harmonic mean of s_d, as noted by Charlesworth (1996).

This role of the harmonic mean, which is influenced more by smaller than larger values, is readily explained intuitively. Mutations of large deleterious effect are driven to low mutation-selection frequencies, while mutations of smaller effect rise to higher equilibrium frequencies. The mean fitness effect of deleterious mutations at equilibrium therefore must be closer to that for mild than strong deleterious new mutations.