Rates of Spontaneous Mutation

John W. Drake,* Brian Charlesworth,† Deborah Charlesworth† and James F. Crow‡

*Laboratory of Molecular Genetics, National Institute of Environmental Health Sciences, Research Triangle Park,
North Carolina 27709-2233, †Institute of Cell, Animal and Population Biology, University of Edinburgh,
Edinburgh EH9 3JT, Scotland, United Kingdom, and ‡Genetics Department, University of Wisconsin, Madison, Wisconsin 53706

ABSTRACT

Rates of spontaneous mutation per genome as measured in the laboratory are remarkably similar within broad groups of organisms but differ strikingly among groups. Mutation rates in RNA viruses, whose genomes contain *ca.* 10^4 bases, are roughly 1 per genome per replication for lytic viruses and roughly 0.1 per genome per replication for retroviruses and a retrotransposon. Mutation rates in microbes with DNA-based chromosomes are close to $1/300$ per genome per replication; in this group, therefore, rates per base pair vary inversely and hugely as genome sizes vary from 6×10^3 to 4×10^7 bases or base pairs. Mutation rates in higher eukaryotes are roughly 0.1–100 per genome per sexual generation but are currently indistinguishable from $1/300$ per cell division per *effective* genome (which excludes the fraction of the genome in which most mutations are neutral). It is now possible to specify some of the evolutionary forces that shape these diverse mutation rates.

RATES of spontaneous mutation per replication per measured target vary by many orders of magnitude depending on the mutational target size (from 1 to $>10^{10}$ b, where b stands for base or base pair as appropriate), the average mutability per b (from 10^{-4} to 10^{-11} per b per replication), and the specific mutability of a particular b (which can vary by $>10^4$-fold). A mutation rate comprises all kinds of mutations in a mutational target: base pair substitutions, base additions and deletions (often producing frameshifting in exons), and larger or more complex changes. Attempts to detect order in these mutation rates have revealed certain underlying patterns. We describe these patterns, note some of their consequences, and consider their evolutionary origins.

Among the mutations that affect a typical gene, different kinds produce different impacts. A very few are at least momentarily adaptive on an evolutionary scale. Many are deleterious. Some are neutral, that is, they produce no effect strong enough to permit selection for or against; a mutation that is deleterious or advantageous in a large population may be neutral in a small population, where random drift outweighs selection coefficients. The impact of mutation is quite different in different DNA sequences. It is maximal in a conventional gene or exon, and at least transitorily less in a gene whose function is required rarely or is redundant. If adaptive mutations are rare, as seems to be the case, then rates of DNA sequence evolution are driven mainly by mutation and random drift, as Kimura (1983a) has argued. In this case, the proportion of neutral mutations at a site or locus is the ratio of its rate of evolution to

that of a region that can be considered neutral, such as a pseudogene. Most newly arisen mutations in functional genes are deleterious, but the fraction may approach zero for spacer DNAs such as introns and intergenic regions. Of course, some protein evolution certainly results from favorable mutations, and to this extent the neutral fraction is overestimated.

The existence of strong taxonomic patterns of mutability implies that genomic mutation rates are close to an evolutionary equilibrium whose driving forces we consider here. The evolution of those rates is likely to reflect their average effect over long periods, but this effect is likely to have been insignificant for much of the spacer DNA.

THE MAGNITUDES OF MUTATION RATES

Terminology: Table 1 describes the abbreviations and parameters we will use to describe the mutation process. Note that the effective genome size G_e is similar to the total genome size G in microbes, whereas $G_e < G$ in higher eukaryotes. The most reliable estimates of mutational parameters come from microbes whose genes are encoded by DNA ("DNA-based microbes"); as we discuss, there are considerable uncertainties attached to estimates from RNA viruses and higher eukaryotes.

Mutation rates in lytic RNA viruses: Few investigators of the genetics of RNA viruses have focused specifically on mutation *rates*, although mutant *frequencies* are often noted to be high compared with those observed in microbes with DNA chromosomes. In a recent survey (Drake 1993a), most of the mutation rates that could be calculated were necessarily based on results obtained with very small and thus potentially unrepresentative mutational targets, and contained other experimental and calculational uncertainties. These uncertainties in-

TABLE 1

Parameters used in describing the mutation process

Parameter	Definition
Abbreviations	
b	Base or base pair
kb	Kilobase or kilobase pair
Parameters	
C	The variously calculated ratio of total mutations to observed mutations
f	Mutant fraction in a population
G	Genome size in bases or base pairs (haploid unless otherwise indicated)
G_e	Effective genome size (portion in which most mutations are deleterious)
h	Dominance coefficient associated with a typical detrimental mutation
ΔM	Rate of decline per generation of the mean of a quantitative fitness trait
N	Population size
$\bar{s}$	Mean reduction in trait value of a single mutation when homozygous
$\tilde{s}$	Mean selection coefficient against a heterozygous deleterious mutation
U	Rate of deleterious mutations per genome per sexual generation
U_c	U for a chromosome
U_d	U for the diploid genome
U_h	U for the haploid genome
ΔV	Increase in variance of a quantitative fitness trait in parallel lines
δ	Reduction in mean fitness of inbred versus randomly mated diploids
μ_b	Mutation rate per base pair per replication
μ_g	Mutation rate per genome per replication
μ_{eg}	Mutation rate per effective genome per replication
μ_{egs}	Mutation rate per effective genome per sexual generation
μ_{lin}	Mutation rate under conditions of linear replication in RNA viruses
μ_{bin}	Mutation rate under conditions of binary replication in RNA viruses
μ_m	$(\mu_{lin} + \mu_{bin})/2$ for RNA viruses

cluded lack of information about what proportion of lytic virus replication is linear (repeated copying of the same template) and what is binary (as in most DNA replication), as well as lack of information about the relative contributions of transcription and reverse transcription to retroviral mutation rates. (J.W.D. wishes to correct a typographical error in his 1993a report: the minus signs do not belong in Equations 2 and 3.)

Many (but not all) of the mutation-rate calculations for these viruses were performed by transforming a mutant frequency f into a mutation rate μ, where f was measured in large populations that had accumulated mutants in the putative absence of selection. For linear replication, $\mu_{lin} = f$ regardless of the extent of growth. For binary replication, $\mu_{bin} = (f - f_0)/\ln(N/N_0)$ where N_0 is the initial and N is the final population size. [This holds for $N_0 > 1/\mu_{bin}$; for $N_0 < 1/\mu_{bin}$, $\mu_{bin} = f/\ln(N\mu_{bin})$.] Because the relative numbers of binary and linear replications are unknown, Drake (1993a) simply averaged μ_{lin} and μ_{bin} to obtain μ_m. μ_{lin} was at most about an order of magnitude greater than μ_{bin} so that μ_m was a little less than $\mu_{lin}/2$ and about sixfold larger than μ_{bin} (range 2.2- to 9.6-fold). In addition to these uncertainties, the relative fidelities of binary and linear RNA replication are yet to be determined, and differences of a few-fold would not be surprising. In addition, the calculational

uncertainties were roughly similar in magnitude to the experimental uncertainties. The results from Drake (1993a) are summarized in Table 2. For the lytic RNA viruses, $\mu_g \approx 1$ but with considerable scatter. Values of $\mu_g > 2$ are likely to be overestimates because such values would tend to extinguish the species.

Because a lytic virus replicates repeatedly in each infective cycle, an infected cell yields virus carrying several new mutations per particle. Most of these will be deleterious. The high mutation rate in these viruses may contribute strongly to their characteristic low specific infectivities (infectious particles per physical particle). (Another contributor to low specific activity is the inherent lability of the RNA backbone.) Such viral populations are ex-

TABLE 2

Mutation rates per genome per replication in lytic RNA viruses

Virus	Median μ_g
Bacteriophage Qβ	6.5
Poliovirus	0.8
Vesicular stomatitis virus	3.5
Influenza A	≥ 1.0

Values are medians from Drake (1993a).

tremely vulnerable to increased mutation rates, even a three-fold increase leading to extinction (Holland *et al.* 1990).

In addition to the entries in Table 2, two reports have appeared in which a mutational target of foreign origin was inserted into a lytic RNA virus that was then passaged extensively and eventually screened for accumulated mutations. In the first case (Kearney *et al.* 1993), the target resided in tobacco mosaic virus which was serially passed through plants ten times, each passage expanding an inoculum of $\sim 10^3$ infective units (iu) to $\sim 10^{15}$ iu. After the final passage, the target sequence was reverse-transcribed and amplified by the polymerase chain reaction (PCR) from a number of isolates and then sequenced. The mutant frequencies were $26/16158$ b sequenced (16.1×10^{-4}) after passage, and $8/8208 = 9.7 \times 10^{-4}$ before passage, the latter reasonably attributable to reverse transcription (RT) and PCR errors. While this difference is not significant, if taken at face value it yields a net $f = \mu_{lin} \approx 6.3 \times 10^{-4}$ per b. If the population is considered to have expanded from 10^3 to $10^{15 \times 10}$ iu, $\mu_{bin} \approx 2 \times 10^{-6}$. Then $\mu_m \approx 3 \times 10^{-4}$ per b and, for $G = 6395$ b (Goelet *et al.* 1982), $\mu_g \approx 2$. This value is typical of lytic RNA viruses. However, because $\mu_{lin}/\mu_{bin} > 300$, this calculation is not robust; μ_g could approach 0.01 if binary replications predominated. Another confounding factor is the possibility of bottlenecks. A target size of about 200 b, an inoculum of about 1000 and a momentary f of 10^{-4} would ensure the transmission of about 20 pre-existing mutants at passage. However, if a small fraction of the inoculum contributed heavily to the whole-plant yield, bottlenecks could still occur and the mutation rate would be underestimated.

In the second case (Schnell *et al.* 1996), the target resided in vesicular stomatitis virus (VSV) which was serially passaged through cultured cells 15 times. Each inoculum of about 10^5 plaque-forming units (pfu) (applied to about 10^7 cells) was expanded to about 10^{11} pfu. The target from six isolates was then reverse-transcribed, PCR-amplified, and sequenced. The mutant frequency was $2/2400$ b with no estimate of the contribution of RT and PCR errors. Here $f = \mu_{lin} \approx 8.3 \times 10^{-4}$ and, with $N_0 = 10^6$ and $N = 10^{15 \times 6}$, $\mu_{bin} \approx 4.3 \times 10^{-6}$; then $\mu_m \approx 4.2 \times 10^{-4}$ per b and $\mu_g \approx 4.7$, a result indistinguishable from the values of 2.8 and 4.3 listed in Drake (1993a). However, because $\mu_{lin}/\mu_{bin} \approx 200$, the calculation is again not robust and μ_g could approach 0.05 if binary replications predominated. Here, a target size of 400 b, an inoculum of about 10^5 and an f of as little as 10^{-4} would ensure the transmission of about 4000 pre-existing mutants at passage, thus preventing bottlenecks. A more arcane possibility is that the target sequence, a bacterial gene encoding chloramphenicol acetyltransferase, provided an unexpected selective advantage when functional; this could be easily measured. In the end, however, a deeper understanding of these numbers will require much more analysis of the relative number and order of linear and binary replication events, including the supra-binary component arising from multiple cell cycles per passage.

Mutation rates in retro-elements: In contrast to the lytic RNA viruses, a retrovirus or retrotransposon chromosome replicates precisely three times per infective cycle. Transcription by the host RNA polymerase produces an RNA genome. Reverse transcriptase then catalyzes two replications to generate a DNA-based chromosome that integrates into the host chromosome (of a different cell in the case of a packaged retrovirus, or of the same cell in the case of a retrotransposon) and thereafter assumes a far lower mutation rate. The resulting mutant frequency is the sum of the mutation rates of the three steps, whose magnitudes have not yet been factored. Table 3 lists those rates described in Drake (1993a) which were based on large mutational targets, together with several measurements reported since 1993, including one for a long-terminal-repeat retrotransposon. These retro-element rates are roughly an order of magnitude lower than the RNA-virus rates listed in Table 2. Because of the large mutational target sizes employed, the rate differences among these viruses may be real. (The rate does not correlate with the retroviral or artificial origin of the mutational target sequence.) Compared to the lytic viruses, the retroviral mutation rates may not appreciably reduce specific infectivity. Spleen necrosis virus is slightly more resistant to increased mutation rates than are lytic RNA viruses, being obliterated only after a roughly 13-fold increase (Pathak and Temin 1992).

Mutation rates in DNA-based microbes: Rates of spontaneous mutation in this class of organisms were last surveyed in Drake (1991) and are summarized in Table 4 using a few updated values for genome sizes. Unlike the experimental and theoretical limits to the accuracy of the RNA-virus values, the DNA-microbe values were determined in well-studied systems using robust calculations, and the individual values are likely to be accurate to within two-fold. Table 4 shows that μ_b and G vary inversely and smoothly over nearly four orders of magnitude while μ_g remains constant. Given the paucity of general, constant values in evolutionary processes, this particular constant is strikingly robust.

Heat promotes a variety of base-loss and base-modification reactions and can be strongly mutagenic. The archeon *Sulfobolus acidocaldarius* growing at 75° produces *pyrE* and *pyrF* mutations at 2.8 ± 0.7 and 1.5 ± 0.6 per 10^7 cell divisions, respectively (Jacobs and Grogan 1997). Although genome size, mutational target size and efficiency of mutation detection are not yet measured in this system, *pyrE* genes contain 600–720 b and *pyrF* genes contain 700–1200 b in several bacteria and eukaryotes, and $G = (2–3) \times 10^6$ for related bacteria (D. W. Grogan, personal communication). Using a typical C value of 3.12 to correct for the efficiency of mutation detection (Drake 1991), $\mu_g = 0.0005–0.005$; using the

TABLE 3

Mutation rates per genome per replication in retroelements

Element	Target (b)	μ_g	References
Retroviruses			
Spleen necrosis virus	*lacZ*α (288)	0.04[a]	Drake (1993a)
Murine leukemia virus	MLV (1380)	≥0.06	Drake (1993a)
	tk (≥1128)	0.26[b]	Parthasarathi *et al.* (1995)
Rous sarcoma virus	RSV (1125)	0.43	Drake (1993a)
Bovine leukemia virus	*lacZ*α (288)	0.027[c]	Mansky and Temin (1994)
Human immunodeficiency virus type 1	*lacZ*α (288)	0.16[d]	Mansky and Temin (1995)
	HIV-1 (9181)	0.22[e]	
Retrotransposon			
S. cerevisiae Ty*1*	Ty*1* (173,043)	0.11[f]	Gabriel *et al.* (1996)

[a] The data underlying this value are in Table 1 of Pathak and Temin (1990). That table described six mutations in the VP212-9 clone and added them to eight; in reality, there were seven mutations, which include F92 (described in their Table 2) (H. M. Temin, personal communication). The total of 37 mutations is correct.

[b] Of 244 tk^- mutants arising at $f = 0.0882$, 114 were of large extent and arose in a target of 2620 b and 130 were small and arose in a target of 1128 b. Among 49 sequenced small mutants, 28 contained non-BPS mutations and three contained nonsense mutations (B. D. Preston, personal communication). Following Drake (1993a), μ_b (large) $= (0.0882/3)(114/244)/2620 = 0.524277 \times 10^{-5}$ and μ_b (small) $= (0.0882/3)(130/244)[28 + 3 \times 64/3]/49]/1128 = 2.607 \times 10^{-5}$ and μ_b (total) $= 3.13 \times 10^{-5}$. $G = 8372.5$ (Drake 1993a), so that $\mu_b = 0.262$.

[c] Of 11 mutations in 18,009 288-b *lacZ*α targets, eight were other than base pair substitutions and two were chain-terminating base pair substitutions. Thus, following Drake (1993a), $f = 6.1 \times 10^{-4}$, $C = [8 + (2 \times 64)/3]/11 = 4.6$, and $\mu_b = Cf = 3.3 \times 10^{-6}$. $G = 8500$ (Sagata *et al.* 1985).

[d] Of 69 mutations in 15,424 288-b *lacZ*α targets, 24 were non-PBS and 10 were nonsense mutations. Thus, following Drake (1993a), $f = 4.5 \times 10^{-3}$, $C = [24 + (10 \times 64)/3]/69 = 3.4$, and $\mu_b = Cf = 1.8 \times 10^{-5}$. $G = 9181$ (Ratner *et al.* 1985).

[e] Using a combination of direct sequencing and single-strand conformational polymorphism analysis, and assuming an overall efficiency of mutation detection of 0.8, a total of 12 mutants were detected among 630,000 base replications (B. D. Preston, personal communication). Thus, $\mu_b = 12/(630,000 \times 0.8) = 2.38 \times 10^{-5}$ and $\mu_g = 0.219$.

[f] The 13 base pair substitutions appear to have comprised only 10 mutational events; $f = \mu = 10/173043$ per b and $G = 5918$.

mid-range values for the above parameters, $\mu_g \approx 0.002$. Thus, although careful measurements remain to be performed in this system, the magnitude of μ_g seems likely to be conserved even in a potentially hypermutagenic environment.

As noted previously, RNA-virus and retrovirus popula-

TABLE 4

Mutation rates per genome per replication in microbes with DNA chromosomes

Organism	G	μ_b	μ_g
Bacteriophage M13	6.4×10^3	7.2×10^{-7}	0.0046
Bacteriophage λ	4.9×10^4	7.7×10^{-8}	0.0038
Bacteriophages T2 and T4	1.7×10^5	2.4×10^{-8}	0.0040
Escherichia coli	4.6×10^6	5.4×10^{-10}	0.0025
Saccharomyces cerevisiae	1.2×10^7	2.2×10^{-10}	0.0027
Neurospora crassa	4.2×10^7	7.2×10^{-11}	0.0030
Mean			0.0034

Values are from Drake (1991), excluding the two well defined outliers and updated slightly as described in the text. The updated G values are 168.8 kb for phage T4 (Kutter *et al.* 1994), 4639 kb for *E. coli* (Blattner *et al.* 1997) and 12068 kb for *S. cerevisiae* (Goffeau *et al.* 1996).

tions are likely to be extinguished when their mutation rates are increased to a few-fold over 1. To be similarly jeopardized, the microbes in Table 4 would have to experience mutation-rate increases on the order of 10^3-fold. However, they are to some extent buffered against immediate extinction in two ways. First, a substantial fraction of their genes are only infrequently required for growth, particularly under laboratory conditions. Second, diploidy, when it occurs, will protect for a while against the effects of recessive mutations. As described in Drake (1991), *Escherichia coli* can survive $\mu_g \approx 10$ for at least 10 generations (although such cultures contain many dead cells), and *Saccharomyces cerevisiae* can survive $\mu_g \approx 60$ for at least nine generations while diploid, although the haploid segregants are inviable; haploids can survive $\mu_g \approx 2$, although most cells grow poorly.

The *E. coli* F plasmid ordinarily replicates in step with the host chromosome, uses most of the same enzymes, and has the same μ_b as the host (Willetts and Skurray 1987; Drake 1991). Kunz and Glickman (1983) and Christensen *et al.* (1985) reported that the F mutation rate increases during conjugation, and Taddei *et al.* (1995) suggested that this might be an example of the μ_g rule implied in Table 4. In this respect F would resemble

phage λ, which displays the host mutation rate as a prophage but the appropriately higher rate when replicating lytically. Unfortunately, while the data of Kunz and Glickman (1983) suggest that the spectrum changes markedly during conjugation, neither F study provides the mutational spectrum needed to calculate μ_b. Using either the calculations of the authors or calculations based on the methods described in Drake (1991) and noting that base pair substitutions seem to occur preferentially during conjugation (Kunz and Glickman 1983), the conjugational F mutation rate appears to be roughly five-fold to perhaps 20-fold higher than the standard rate. The *E. coli* genome is about 47-fold larger than the F genome (Willetts and Skurray 1987). Thus, additional measurements are needed to determine whether F follows this μ_g rule.

Two predictions: A plot of log μ_b versus log G (Drake 1991) reveals a gap between the viral and the cellular entries. The 578-kb genome of *Mycoplasma genitalium* (Peterson *et al.* 1995) falls at the midpoint of this gap. An interpolation from Table 4 predicts $\mu_b = 5.9 \times 10^9$ for this organism.

Mutation rates in bacteriophage T4 and herpes simplex virus type 1 (HSV) display an intriguing relationship which suggests that at least one strain of HSV may harbor a mutator mutation. These two viruses have similar genome sizes and modes of DNA replication. In HSV stocks grown from small inoculum to $N = 10^8$–10^9 iu, the frequency of herpesvirus *tk* mutants is about 6.2×10^4 (Hall *et al.* 1984). The *tk* gene has about 1150 b (McKnight 1980) and $G = 152{,}260$ b (McGeoch *et al.* 1988). No suitable mutational spectrum is available, so C must be guessed by taking the average of the values listed in Drake (1991); then $C = 3.12$, $\mu_b \approx 1.7 \times 10^7$ and $\mu_g \approx 0.026$. This μ_g is roughly eight-fold higher than the values in Table 4. Consider next the antimutator mutations that arise in the DNA polymerase gene of phage T4. These reduce the rates of only certain pathways, while increasing the rates of others; overall, they do *not* reduce μ_b (Drake 1993b). A strong, general antimutator is probably difficult or impossible to obtain in one or a few mutational steps (Drake 1993b), the exception being by the reversal of a mutator mutation that itself arose by a single mutation. HSV DNA polymerase mutants selected for resistance to phosphonoacetic acid (PPA) are sometimes antimutators. These reduce the frequency of *tk* mutants of presumably many kinds (Hall *et al.* 1984), and therefore presumably *do* reduce μ_b; the reduction is roughly 45-fold, giving $\mu_g \approx 0.0006$ (about five-fold lower than than the values in Table 4). In contrast, wild-type phage T4 is resistant to PPA; however, T4 DNA-polymerase mutator mutations are sensitive to PPA. When selection is then applied for PPA resistance in these mutator strains, the result is polymerase antimutator mutations that negate the mutator phenotypes (Reha-Krantz *et al.* 1993). These results suggest that this HSV isolate may carry a naturally

occurring mutator mutation. Results described next reveal that this is a reasonable conjecture.

Microbial hypermutation: Microbial mutation rates can increase over short periods for physiological or regulatory reasons, or more permanently by the action of mutator mutations. In addition, particular portions of the genome can be maintained inherently hypervariable through specific, local mechanisms such as the cassette switching that mediates phase variation in bacterial and other pathogens and mating type in yeasts and fungi (Moxon *et al.* 1994; Sasaki 1994).

Microbial mutation rates can increase physiologically in several ways. Ninio (1991) suggested that errors of transcription, translation, and molecular segregation will create transient mutators which, on his estimate, would contribute modestly to single mutations but strongly to multiple mutations per genome per replication. In *Neurospora crassa* (Auerbach 1959) and phage T4 (Drake 1966; Drake and Ripley 1994) (and probably in all organisms at various times), resting genomes mutate in a time-dependent, replication-independent fashion because they accumulate spontaneous DNA damage that engenders mutations when DNA replication resumes, and that may even alter transcription to produce a mutant phenotype before replication. Starving bacteria also mutate in a time-dependent manner, one that probably involves immediate DNA synthesis (*e.g.*, Foster 1997; Torkelson *et al.* 1997). In addition, DNA damage sometimes elicits the process of translesion bypass, in which DNA primer extension passes an unrepaired lesion. Among microbes, this process can range from fully constitutive, as in phage T4 (Drake and Ripley 1994), to strongly inducible, as in the *E. coli* SOS response (Walker 1984). The SOS response increases mutation rates for roughly one cell generation, even in undamaged parts of the genome (Walker 1984). Because DNA damage is a byproduct of ordinary endogenous processes such as base depurination and deamination, base damage from by-products of oxygen and methyl metabolism and so on, a few cells in any population are SOS-induced at any time. The fraction of evolutionary change driven by such transient hypermutability remains unknown.

For bacteria, chemostats or daily serial transfers constitute alien environments within which rapid and complex adaptation occurs by mutation and selection. Because at least 10 genes can generate mutator mutations, *E. coli* populations generate roughly 10^{-6}–10^{-5} mutator mutants per replication. However, strong mutators are deleterious (Quiñones and Piechocki 1985) and do not accumulate; an ordinary *E. coli* culture accumulates $<10^{-5}$ mutator mutants (Mau *et al.* 1997). On the other hand, mutators can be strongly selected when their frequency and strength are high enough so that they generate more beneficial mutations than do the non-mutators in the same population (Chao and Cox 1983; Mao *et al.*

1997); the deleterious mutators are selected indirectly along with the adaptive mutation.

In contrast to freshly grown laboratory cultures, mutator mutants are found at frequencies that can exceed 10^{-2} among hospital isolates of *E. coli* and *Salmonella enterica* (Jyssum 1960; Gross and Siegel 1981; LeClerc *et al.* 1996), or after extensive serial passage in the laboratory (Sniegowski *et al.* 1997). Mutator mutations are common among commensal as well as pathogenic strains, and may increase mutation rates either strongly or weakly (Matic *et al.* 1997). Thus, the continual adaptations occurring during bacterial invasions of new hosts or culture conditions suffices to enhance the frequency of mutators by at least 1000-fold, although the subsequent deleterious effects of the mutator mutations may prevent total replacement among hosts or serially transferred lines. In order to understand the roles of mutator mutations in both transitory adaptations and long-term evolution, it will be important to determine the frequencies of mutators in natural populations in both stable and strongly fluctuating environments. Theory describing conditions under which mutators can speed adaptation in asexual microbes (Leigh 1970; Taddei *et al.* 1997) encourages such investigations.

Mutation rates in higher eukaryotes based on specific loci: Plants and animals contrast with the organisms in Tables 2–4 in several ways. One striking difference is in the amount of DNA. G is one to several orders of magnitude greater in plants and animals than in microbial eukaryotes. Most of the increase is not in functional genes but rather in introns and inter-genic regions, so that $G_e << G$. A second difference is that higher eukaryotes may display important age and sex effects. As we discuss below, in mammals (and especially in humans), the rate of gene mutation per generation is much higher in males and particularly older males, mainly because of the much larger number of germ-line cell divisions ancestral to a sperm than to an egg. A third difference is that mutation rates in animals (and plants) are often equated with the mutant frequency per gamete (or, occasionally, per diploid). Sometimes, however, mutants appear in clusters that reflect the premeiotic expansion of a single event (Muller 1952; Woodruff *et al.* 1997). Unrecognized clusters are not a problem, because a cluster increases proportionally the probability of finding the mutation. When a cluster is observed, each mutant individual in the cluster should be counted as a mutation when calculating the mutation rate per sexual generation; more complex calculations of within-generation mutation accumulation are difficult because of uncertainties about the topology of normal and mutant cell expansion. A fourth difference is that evolutionary mechanisms for adjusting mutation rates may be quite different in sexual eukaryotes than in rarely sexual microbes because, at least in outbreeding sexual species, the process of meiosis uncouples rate-adjusting mutations from the mutations they engender (Leigh 1970,

1973). Also, sexual reproduction permits the population to rid itself of deleterious mutations more efficiently than is possible in asexual systems (Kimura and Maruyama 1966; Kondrashov 1984, 1988).

In the species we discuss, the data for mutation rates in males are often more extensive and reliable than those for rates in females, so that our calculations frequently must focus on data from males. In all of these species, mutations with small effects tend to go uncounted. Unlike the situation with microbes, where mutational spectra predict the efficiency of detection, the present values are all minimum estimates uncorrected for poorly detected kinds of mutations.

Zea mays: Plants have yielded remarkably few estimates of mutation rates. In plants such as maize where genetic methods are well established, mutation rates are relatively easily estimated by crosses to strains homozygous for mutations causing visible phenotypes. There is wide variation from locus to locus, with a mean of 7.7×10^{-5} and a range of from $<0.1 \times 10^{-5}$ (*waxy*) to 49.2×10^{-5} (*R*) mutations per gamete for eight maize loci (Stadler 1930). There are few comparable data from other plants apart from evidence for lower rates in polyploids (Stadler 1929). Lack of further information impedes attempts to extrapolate to the entire genome.

Caenorhabditis elegans: There are about 8.2 cell divisions ancestral to sperm and about 10.0 ancestral to eggs (Kimble and Ward 1988), so we will use the average, 9.1. Mutation rates and numbers of codons have been determined for five loci. Taking the gene sizes as three times the number of codons plus 100 b for regulatory and splicing sequences, using an average $C = 3.12$ to correct for the efficiency of mutation detection (Drake 1991) and dividing by the 9.1 cell divisions per sexual generation, we obtain $\mu_b(unc\text{-}22) = 1.17 \times 10^{-11}$, $\mu_b(unc\text{-}54) = 1.68 \times 10^{-11}$, $\mu_b(unc\text{-}93) = 52.4 \times 10^{-11}$, $\mu_b(unc\text{-}105) = 4.54 \times 10^{-11}$, and $\mu_b(sup\text{-}10) = 52.6 \times 10^{-11}$ (Greenwald and Horvitz 1980; Karn *et al.* 1983; Eide and Anderson 1985; Benian *et al.* 1989; Levin and Horvitz 1992; Liu *et al.* 1996; C. White and P. Anderson, GenBank accession no. U43891); the mean $\mu_b = 2.25 \times 10^{-10}$. The total genome size $G = 8 \times 10^7$ (Sulston and Brenner 1974). There are about 1.78×10^4 genes (Bird 1995); assuming an average of 10^3 b per gene gives $G_e = 1.78 \times 10^7$. Then $\mu_g = 8 \times 10^7 \times 2.25 \times 10^{-10} = 0.018$ and $\mu_{eg} = 1.78 \times 10^7 \times 2.25 \times 10^{-10} = 0.0040$. These values are posted in Table 5.

Drosophila melanogaster: Schalet (1960) detected 51 visible mutations in 490,118 *X* chromosomes at 13 specific loci, yielding a rate of 8.0×10^{-6} per locus per generation. The fraction of these due to base pair substitutions is unknown; given the evidence that a large fraction of visible mutations in Drosophila are caused by insertions of transposable elements (Finnegan and Fawcett 1986), it is probable that at least half of Schalet's mutations were of this nature. Mukai and Cockerham (1977) enriched the mutation frequency by ac-

TABLE 5

Mutation rates estimated from specific loci in higher eukaryotes

Organism	G	G_e	μ_b	μ_g	μ_{eg}	μ_{egs}
C. elegans	8.0×10^7	1.8×10^7	2.3×10^{-10}	0.018	0.004	0.036
Drosophila	1.7×10^8	1.6×10^7	3.4×10^{-10}	0.058	0.005	0.14
Mouse	2.7×10^9	8.0×10^7	1.8×10^{-10}	0.49	0.014	0.9
Human	3.2×10^9	8.0×10^7	5.0×10^{-11}	0.16	0.004	1.6

These are all likely to be underestimates, because not all kinds of mutations are included, particularly those with minor, inconspicuous effects.

cumulating isozyme mutations in 1000 chromosomes sheltered by heterozygosity in a balanced-lethal system for almost 175 generations. In 1,658,308 locus-generations they found three electrophoretic-mobility (band-shift) mutations and 17 null (band-loss) mutations. However, these strains exhibited a high rate of chromosome breakage, probably because of an active transposon (Yamaguchi and Mukai 1974); it is therefore probably more realistic to ignore the nulls, a procedure also justified by the high average ratio of base pair substitutions to other mutations in microbes (Drake 1991). Mukai and Cockerham (1977) estimated that about 0.3 of all amino acid changes were detectable as band shifts. In addition, only about 2/3 of base pair substitutions change an amino acid. Thus, the mutation rate per locus per generation is $(3/1,658,308)/(0.3)(2/3) = 9.0 \times 10^{-6}$. Averaging the two studies, we take 8.5×10^{-6} as a representative rate. The proteins studied by Mukai and Cockerham (1977) were encoded by an average of 973 b and some regulatory sequences must also have been present, so dividing by 10^3 gives 8.5×10^{-9} mutations per b per generation. The number of cell divisions ancestral to a sperm in Drosophila is about 25 for the young males typically used in laboratory experiments (Lindsley and Tokuyasu 1980; Drost and Lee 1995; J. M. Mason, personal communication), so dividing by 25 gives $\mu_b = 3.4 \times 10^{-10}$. In Drosophila, $G \approx 1.7 \times 10^8$ b (Ashburner 1989). We will take as G_e the amount of DNA in 1.6×10^4 genes, each of length 10^3 b (Bird 1995); this gives $G_e = 1.6 \times 10^7$ b. These and derivative values are given in Table 5.

Mus musculus: The mouse data come from the controls for the extensive radiation experiments performed at Oak Ridge, Harwell and Neuherberg, and summarized by Russell and Russell (1996). A total of 1,485,036 progeny harbored 69 visible mutations at seven loci for a rate of 6.6×10^{-6} per locus per generation. In addition to the complete mutations, about 4.8×10^{-5} mosaic mutations were detected at five loci; these mutations tended to produce about 50% germ-line mosaicism, so that the adjusted mosaic rate is $(4.8 \times 10^{-5})(0.5)/5 = 4.8 \times 10^{-6}$. Thus, the total mutation rate was about 1.1×10^{-5} per locus per generation. Assuming 10^3 b per locus, we obtain 1.1×10^{-8} mutations per b per generation.

Finally, dividing by 62, the estimated number of cell divisions prior to a sperm (Drost and Lee 1995), gives $\mu_b = 1.8 \times 10^{-10}$ mutations per b per cell division. Taking $G = 2.7 \times 10^9$ b (Laird 1971) and G_e as the amount of DNA required for 8×10^4 genes (Bird 1995) of length 10^3 bp generates the additional values listed in Table 5.

It is surprising that germ-line mosaics were responsible for almost 40% of the total rate. These appear to arise either as mutations that occur in DNA replications directly before or after meiosis in the male parent ("after" denoting the first zygotic replication) or in a single strand of nonreplicating DNA (which might include mutations arising during DNA recombination or repair). Thus, almost as many mutations may occur in or between one or two special DNA replications as occur in the other 60. This possibility presents a major challenge to both experimentalists and theoreticians (Woodruff *et al.* 1997).

The mutation rate of evolutionary importance is of course the average over the two sexes. The estimated rates per generation in male and female mice are not very different, but the female value is based on very small numbers and is complicated by a large cluster. Adding the mosaic rate of 4.8×10^{-6} to the female rate of 1.6×10^{-6} gives a rate of 6.4×10^{-6}, about half the male rate. Alternatively, we note that the estimated number of cell divisions prior to the gamete is 25 in females and 62 in males, so the female rate is $25/62 = 0.40$. Averaging these, the female rate is 0.45 of the male rate. The murine male μ_{egs} is 0.55, so the average of the two sexes is about 0.4. A similar result obtains in humans (see below).

Homo sapiens: The human data are less reliable than the *C. elegans*, Drosophila and mouse data. A number of dominant-mutation rates have been inferred from the frequency of affected children of normal parents, and sometimes confirmed by equilibrium estimates for those dominants with severe effects. These values range from 10^{-4} to 10^{-6}, with a rough average of 10^{-5} (Vogel and Motulsky 1997). For genes of size 10^3 b, this corresponds to a rate of 10^{-8} per b per generation. An estimate based on specific changes in the hemoglobin molecule gave 0.74×10^{-8} per b per generation (Vogel and

Motulsky 1997), but this is clearly an underestimate because other kinds of changes are not included. A third, quite independent estimate is based on rates of evolution of pseudogenes in human ancestry, which are likely to be identical to mutation rates (Kimura 1983a). This gives about 2×10^{-8} per b per generation (Crow 1993, 1995). We shall take 10^{-8} as a representative value. However, because the overwhelming majority of human mutations occur in males (see below), the male rate must be about twice the average rate, or 2×10^{-8}. The number of cell divisions prior to sperm formation in a male of age 30 is about 400 (Drost and Lee 1995; Vogel and Motulsky 1997). Thus, $\mu_b \approx 2 \times 10^{-8}/400 = 5 \times 10^{-11}$. For 8×10^7 genes (Bird 1995) of average size 10^3 b, $\mu_{eg} \approx 0.004$ and $\mu_{egs} \approx 1.6$.

An alternative method for estimating μ_{egs} has been proposed by Kondrashov and Crow (1993) based on the idea that purely neutral sequences such as pseudogenes can be used as a benchmark to identify sites which show clear evidence of selective constraints. If the abundance of such sites can be determined in this way, the effective genome size and its mutation rate can be estimated purely from rates of DNA sequence evolution. This method has yet to be applied to large quantities of sequence data. For hemoglobin genes, about 15% of bases seem to be under the effective control of selection, which may be about average for genes encoding proteins; for a more sophisticated treatment, see Kimura (1983b).

With 6.4×10^9 base pairs in the diploid genome, a mutation rate of 10^{-8} means that a zygote has 64 new mutations. It is hard to image that so many new deleterious mutations each generation is compatible with life, even with an efficient mechanism for mutation removal. Thus, the great majority of mutations in the noncoding DNA must be neutral.

Effects of sex and age in humans: Data for female mutation rates are less numerous and less reliable than data for male rates. For human base pair substitutions, there is an enormously greater mutation rate in males than in females, along with a strong paternal age effect. Older males have a higher rate than younger males, and the increase is greater than if mutation were simply cell division-dependent (Crow 1993, 1997), but the component of replication fidelity that decreases in older men remains unknown. Mutations also accumulate with cell divisions in somatic tissue (Akiyama *et al.* 1995), although whether the rate is proportional to number of divisions is not known.

The enormous difference in human male and female mutation rates is well documented for those few loci for which there are adequate data. For three conditions, Apert's Syndrome, multiple endocrine neoplasia type B (MEN2B), and type A (MEN2A), a total of 92 new mutations have been reported in which, by linkage analysis, it is possible to determine the parent of origin. Strikingly, all 92 were paternal. These are all base-substi-

tution mutations. Apert's Syndrome has also been studied for paternal age effect and, as expected, shows a large increase with paternal age (Crow 1997). The fact that so many of these mutations are at CpG sites offers some support to those who argue that something associated with methylation is responsible for the high male rate (Sapienza 1994; McVean and Hurst 1997).

In contrast, some mutations are not strongly associated with paternal age. Two examples are neurofibromatosis and Duchenne muscular dystrophy. In both of these diseases most of the mutations are small deletions and other cytogenetic changes in enormous genes. Thus, the generalization seems to be that base substitutions are replication-dependent but cytogenetic changes are not. Various human diseases show a continuum from very strong paternal age effect to very little (Risch *et al.* 1987), presumably reflecting the relative proportions of base substitutions and cytogenetic changes.

Effect of hemizygosity: Deleterious mutations at sex-linked loci are strongly expressed in the hemizygous state in the heterogametic sex and are thus subject to stronger counter-selection than are deleterious mutations in autosomal genes (except when strongly dominant) (Haldane 1927). As a result, there is stronger selection pressure to reduce mutation rates at X-linked loci than at autosomal loci (McVean and Hurst 1997). Data on mammalian DNA sequence evolution suggest that the X chromosome may indeed experience a lower mutation rate than the autosomes (McVean and Hurst 1997), although this may be confounded with a higher mutation rate in males than in females, especially in long-lived mammals.

Mutational hot spotting: Some of the best understood human mutations arise in the gene for achondroplasia, which would seem to be a good source for a mutation rate estimate. The average mutation rate for the phenotype, determined directly in several studies and substantiated by indirect calculations, is 10^{-5} (Vogel and Motulsky 1997). However, molecular analysis (Shiang *et al.* 1994) revealed that 15 of 16 mutations were GGG → AGG and the other was GGG → CGG at the same codon, replacing glycine with arginine. Thus, the entire observed mutation rate appears to come from one codon. Similar CpG hotspots were responsible for all the mutations causing Apert's Syndrome. Although the data are scanty, these two examples suggest that a major fraction of human gene mutations is due to mutational hot spots, as is also typical in microbes (*e.g.*, Benzer 1961). We badly need more data on per-locus mutation rates accompanied by molecular analyses showing the mutant sites and the parent of origin.

Somatic versus germinal mutation rates: The (mammalian male) germ-line rate may be lower than the somatic rate. Using mice bearing a chromosomal mutation-reporter target, the typical somatic-cell mutant *frequency* was found to be about 1.7×10^{-5} in a variety of tissues, but about 0.6×10^{-5} in sperm preparations (Kohler *et al.*

1991). (This three-fold difference is probably an underestimate, because the germ-cell preparations used in these experiments were probably contaminated with somatic cells.) Because the number of cell divisions antecedent to these tissues was not notably lower in germ cells than in somatic cells, this result suggests a mutation *rate* per cell division (or unit of time) at least three times lower in germ cells than in somatic cells. In turn, this conclusion suggests that evolutionary pressures on mutation rates focus primarily upon the germ line with the soma being subject to less pressure, perhaps because of diploidy (Orr 1995).

Mutation rates for deleterious alleles from mutation-accumulation experiments: Measurements of mutation rates based on specific loci offer the potential of scoring all mutational events in a molecularly well defined target. This has been achieved in several microbial systems, but as yet only imperfectly in higher eukaryotes. A disadvantage of the specific-locus method is that only one or a few, possibly unrepresentative, genes may have been examined. An alternative approach is to accumulate mutations with deleterious fitness effects over many generations. While this method ignores mutations without effects on fitness, it can screen rather large fractions of the genome. Mutations with very small fitness effects (of the order of 10^{-3} or less) are not likely to contribute to mutation-rate estimates obtained in this way, unless they are improbably numerous; we therefore distinguish these estimates conceptually from the mutation rates per effective genome discussed above. Given that the specific-locus results for higher organisms depend largely on the detection of mutations with major phenotypic effects, the above estimates of the effective genome mutation rates should in practice be quite similar to the deleterious mutation rate estimates.

Mutation-accumulation methods: This procedure combines Muller's (1928) use of marked crossover-suppressing chromosomes to detect mutations anywhere along a chromosome that has been sequestered for several generations, and Bateman's (1959) proposal to use the variance between replicates to estimate the mutation rate and average effect. The approach was refined by Mukai and co-workers (Mukai 1964; Mukai *et al.* 1972), and we shall refer to it as the Bateman-Muller-Mukai method.

The basic idea is as follows. A set of initially genetically identical lines is established from an isogenic base. The lines are maintained independently and mutations are allowed to accumulate. Because the mating system assures that the mutation-accumulating chromosomes reside only in heterozygous males and a single male is used as the parent in each generation, selection is minimized. Assuming that the number of mutations per line after some number of generations is a random variable, different lines will accumulate different numbers of mutations. Thus, the variance among lines for a quantitative fitness trait such as viability will increase over time. Given the deleterious nature of most mutations, the mean

value of the fitness trait is expected to decline with time. Let the mean number of deleterious mutations that arise per generation be U (U_c for the rate for a particular chromosome, U_h for the haploid genomic rate, U_d for the diploid genomic rate), and the mean reduction in trait value of a single mutation when homozygous (relative to a value of 1 for wild-type) be s. (s is a weighted mean, in which the effects of mutations at individual loci are weighted by the mutation rates at the loci.) If mutational effects are additive across loci, the rates of decline in overall mean fitness (ΔM) and increase in variance among the lines (ΔV) are given by

$$\Delta M = U_c\,s \tag{1a}$$

and

$$\Delta V = U_c(s^2 + V_s), \tag{1b}$$

where V_s is the variance among sites in the effects of mutations (again with a weighting by the mutation rates). These yield the expressions

$$U_c \geq \Delta M^2 / \Delta V \tag{2a}$$

and

$$s \leq \Delta V / \Delta M. \tag{2b}$$

These are *equalities* only if all mutations have equal effects on the trait. However, given reliable estimates of the decline in mean and increase in variance, a useful lower bound on the deleterious mutation rate per genome can be obtained from Expression 2a. The true values of the parameters may differ greatly from these bounds. For example, if the mutation effects have an exponential distribution, then the estimate of U_c is doubled and that of s is halved (Crow and Simmons 1983). Under specific assumptions concerning the shape of the distribution of mutational effects, such as a gamma distribution, maximum likelihood methods can be used to estimate the parameters of the distribution and the value of U_c (Keightley 1994, 1996). In principle, these methods should provide more accurate estimates of U, s and V_s than the simpler methods of Mukai *et al.* (1972) provided that the assumptions of the statistical model are met.

There are two difficulties in applying this method of estimating U in species other than Drosophila. The first is the problem of preventing the operation of selection, which obviously opposes the accumulation of deleterious mutations. This can be fairly easily achieved by maintaining each line with minimal effective population sizes, because a mutation is effectively neutral when the product of effective population size and selection coefficient is less than one (Fisher 1930). In clonal or selfing organisms, propagation of each line through a single individual each generation ensures that even highly deleterious mutations will behave like neutral alleles (Keightley and Caballero 1997). Even for diploids with separate sexes, only strongly deleterious al-

TABLE 6

Mutation rates to detrimental alleles from mutation-accumulation experiments

Organism	Trail studied	U_h	Reference
E. coli	Rate of reproduction	≥ 0.0002	Kibota and Lynch (1996)
C. elegans	Lifetime reproductive success	≥ 0.0006	Keightley and Caballero (1997)
A. thaliana	Lifetime reproductive success	≥ 0.1	S. Schultz and J. A. Willis, personal communication
D. melanogaster	Egg-to-adult viability	≥ 0.35[a]	Mukai (1964)
		≥ 0.42[a]	Mukai *et al.* (1972)
		≥ 0.15[a]	Ohnishi (1977)

[a] The genomic rate is estimated by assuming that chromosome *2* constitutes 40% of the genome.

leles will be efficiently eliminated by selection, because an effective population size of two for each line is possible with full-sib mating. Such experiments therefore provide estimates of the rate of mutation to most detrimental alleles, *i.e.*, mutations that reduce fitness by less than about 50% when homozygous.

The main difficulty with this method is that lines accumulating large numbers of mutations become increasingly vulnerable to loss due to low fitness, introducing a downward bias in both ΔM and ΔV. In *D. melanogaster*, the use of marked balancer chromosomes means that mutations can accumulate on a single autosome that is propagated through a single heterozygous male in each line (Mukai 1964); then the effective population size is one-half. Given the considerable recessivity of most strongly deleterious mutations (Crow and Simmons 1983), the risk of loss of a line is greatly reduced by this procedure; in fact, in Mukai's experiments only a very small fraction of the singly-mated males were sterile, and those cultures were replaced by sibs so that no lines were lost (Mukai *et al.* 1972). The chromosomes accumulating mutations can be made homozygous when desired, and their effects on fitness components assayed.

The mutation rate for recessive lethals can also be estimated by the balancer chromosome technique, and is about 0.01 per haploid genome per generation in *D. melanogaster* (Crow and Simmons 1983). This seems to be only a small fraction of the total mutation rate to deleterious alleles (see Table 6). More limited information on *C. elegans* suggests a slightly lower lethal mutation rate of about 0.007 per haploid genome per generation (Rosenbluth *et al.* 1983; Clark *et al.* 1988). However, the upper 95% confidence interval for this estimate overlaps the Drosophila value. In plants, direct estimates can be obtained for nuclear mutations to chlorophyll deficiency (albino and yellow seedlings). These are only a component of all lethal mutations, and so provide an underestimate of the mutation rate to lethal alleles (Klekowski 1988). There is also some information from ferns, in which newly arisen lethals affecting the haploid stage can be detected by testing for inviable spores from individual diploid sporophytes (which must necessarily derive from haploids free from lethals in the

previous generation). Klekowski (1973) found rates of about 0.01 to 0.015 per haploid genome for three fern species, values remarkably similar to the *C. elegans* and Drosophila estimates.

A more serious problem concerns the proper control for the estimation of ΔM, because fitness components are notoriously sensitive to environmental effects (Houle 1992). Thus, mutation-accumulation lines should ideally be measured at the same time as a control that has not had the opportunity to accumulate mutations. An isogenic stock that initially has the same genetic composition as the mutation-accumulation lines does *not* provide a suitable control, even if it is maintained as a random-bred stock with large effective population size as in some experiments (Fernandez and Lopez-Fanjul 1996). This is because most deleterious mutations have only small effects on fitness when heterozygous, and thus can persist in the population for many generations before elimination (Crow and Simmons 1983). The rate of decline in mean fitness of an initially isogenic population as new mutations appear will at first be nearly the same as for the accumulation lines, and will only approach zero after several tens of generations.

This problem of a suitable control can be overcome with organisms that can recover well from freezing, such as *E. coli* and *C. elegans*; suitable methods have become available only recently for Drosophila (Stepkonkus and Caldwell 1993). Mukai and co-workers (Mukai 1964; Mukai *et al.* 1972) and Ohnishi (1977) either used no control or used values from the lines that were presumed to be mutation-free because they retained maximal fitness. These procedures have led to criticism of the values from these experiments (Keightley 1996).

The magnitude of U: Estimates of minimum (Bateman-Muller-Mukai) detrimental mutation rates from mutation-accumulation experiments in several species are summarized in Table 6. As expected from Table 4, the lower bound for *U* in *E. coli* is extremely small, about 6% of the total genomic mutation rate of 0.0034. If the estimated number of cell divisions (36) in the male germ line of Drosophila is multiplied by the *E. coli U* of 0.0002, the resulting Drosophila $U_h = 0.007$. This is much less than the measured estimate of about 0.3. This ratio (0.007:0.3) is roughly the same as the ratio of

E. coli and Drosophila genome sizes (Kibota and Lynch 1996). In Drosophila, however, the ratio of the effective genome size to the total genome size ($G_e{:}G$) is much smaller than in *E. coli.* If only the ratio of effective genome sizes is considered, a large discrepancy remains. In contrast to the Mukai estimates of U_h for Drosophila of 0.15 (Ohnishi 1977) to 0.42 (Mukai *et al.* 1972), U_h for *C. elegans* was estimated to be 0.003 using the fit to a gamma distribution of mutational effects, or 0.0006 using the Mukai method (Keightley and Caballero 1997). Because the genome sizes of *C. elegans* and Drosophila are similar, an estimated two orders of magnitude difference in U_h is disturbing. On the other hand, the estimate for *Arabidopsis thaliana* (0.1) appears to be only slightly lower than for Drosophila, but the confidence intervals on this estimate are very wide (S. Schultz and J. H. Willis, personal communication).

Several possible explanations can be imagined for these discrepancies. One is that the Drosophila estimates are based on measurements of egg-to-adult viability under competitive conditions, whereas the *C. elegans* and Arabidopsis results were for net reproductive output under noncompetitive conditions; differences in fitness are likely to to be more easily detected under harsher conditions (Kondrashov and Houle 1994). Another possibility is that transposable element insertions played a much larger role in causing deleterious mutations in the Drosophila experiments than in the *C. elegans* experiments, where the line involved lacked transposable element activity (Keightley and Caballero 1997). A less palatable explanation is that the Drosophila values for ΔM are gross overestimates caused by adaptive changes in the balancer chromosome against which the mutation-accumulation chromosomes were competed; this would cause an artifactual decline in the mean viability of the mutation-accumulation chromosomes relative to the balancer (Keightley 1996). However, this artifact seems unlikely because the balancer chromosomes were long-established laboratory strains that were likely to be at equilibrium, and in the longest accumulation experiments (Mukai 1969), the viability of the balancer chromosomes would have had to double to account for the observed results. The fact that the (probably minimal) effective genomic mutation rate in Drosophila in Table 5 is far closer to the U estimates in Table 6 for Drosophila than for *C. elegans* further suggests that this putative artefact is improbable.

To avoid the problems of the control for the Drosophila experiments, Garcia-Dorado (1997) used a statistical method for estimating mutational parameters which does not require knowledge of ΔM, but simply fits the observed distribution of mutation-accumulation line values to an assumed form of continuous distribution of mutational effects. This leads to a much smaller estimate of U for the Mukai and Ohnishi experiments ($U_h \approx$ 0.025) than does the Bateman-Muller-Mukai method, although still substantially larger than the *C. elegans*

value. The difficulty with this result is that there is no *a priori* justification for the assumed distribution of mutational effects; for example, there could be one class of mutations with similar but fairly large effects, and another class with much smaller but highly variable effects, as suggested by Keightley (1996). This could have substantial effects on the estimate of *U.* Further research is clearly needed to resolve these uncertainties.

A method that offers a partial solution to the inadequate Drosophila controls has been devised by S. A. Shabalina, L. Y. Yampolsky and A. Kondrashov (personal communication). A large, randomly bred stock is maintained so as to minimize the opportunity for selection on viability and fertility. If this is the case, the mean value of a fitness component should decline at the rate given by Equation 1a, where U_c is replaced by U_d and the selection coefficients are for heterozygotes rather than homozygotes. If a comparable randomly bred stock is maintained under selective conditions that have prevailed for a long time, so that it is at mutation-selection equilibrium, its mean should remain constant except for environmental fluctuations, so that ΔM can be estimated by adjusting for changes in the control. To avoid the possibility of adaptive change in the control stock, the number of generations over which it is maintained may be minimized by keeping it at a low temperature to reduce the number of control generations or using recently developed methods for freezing Drosophila. Their measurements of net fitness under competitive conditions suggested that $\Delta M \approx 0.02$. For Mukai *et al.* (1972) the value was 0.004 per chromosome, which translates into 0.02 per diploid genome. These values are in good agreement, but this is perhaps due to opposing errors. The data of Shabalina *et al.* are for fertility. There is evidence for a more important contribution from fertility than viability to the genetic load for total fitness (Knight and Robertson 1957; Sved 1975; Simmons *et al.* 1978). On the other hand, they measured heterozygous rather than homozygous effects. In any case, the experiments of Shabalina *et al.* support a substantial decline in fitness under circumstances where the control strain had been kept frozen, thus arguing against improvement of the reference population in the Mukai experiments.

Indirect estimates of *U*: Several indirect methods have been proposed for estimating the genomic deleterious mutation rate. For lethal mutations in outcrossing plants, the classical formula for the equilibrium frequency q of a recessive-lethal allele under mutation-selection balance, $q \approx (\mu/s)^{0.5}$ (Haldane 1927), can be applied to the frequencies of nuclear gene-controlled chlorophyll-deficient variants to obtain the mutation rate μ per haploid genome, assuming that the selection coefficient $s = 1$ (Crumpacker 1967; Ohnishi 1982; Klekowski 1988). If lethal mutations are not completely recessive, as is suggested by the Drosophila data (Crow and Simmons 1983), this procedure underesti-

mates the mutation rate because a higher value of μ is needed to compensate for the elimination of heterozygous lethals by selection.

The results for total mutation rates to chlorophyll deficiency were reviewed by Klekowski (1992) for ten species of annual plants from five families. He concluded that rates are surprisingly constant, ranging from $(0.16-0.45) \times 10^{-3}$. There is no apparent relation to the species' DNA content, even though this differed more than 20-fold. The differences in G_e are presumably much less. It is difficult to extrapolate these mutation rates to the whole genome, because the proportion of vital loci contributed by chlorophyll genes is presently unknown.

Estimates can also be obtained for populations whose selfing rates are known, using the equilibrium formulae of Ohta and Cockerham (1974). Estimates for such populations of long-lived mangrove species are $(2.1-5.8) \times 10^{-3}$ (Klekowski and Godfrey 1989; Klekowski *et al.* 1994), about 13-fold higher than those from annuals. This difference would be expected from an increase in mutation rates with age caused by greater numbers of cell divisions before meiosis in old plants; note that in plants, reproductive structures are formed from vegetative tissues, and are not derived from a germ line (Klekowski 1988). This value needs confirmation in other species. Assuming that mutations affecting sporophyte viability also lower male gametophyte viability, an age effect should also decrease pollen viability in old plants. No such effect was found in the one study of which we are aware (Connor and Lanner 1991). However, by analogy with human studies, mutations that accumulate with age are likely to be mainly base substitutions, which are more likely than deletions to pass through the haploid gametophyte (Khush and Rick 1967).

The use of formulae based on the assumption of mutation-selection equilibria has been extended to estimating detrimental mutation rates from the effect of inbreeding on fitness components (Charlesworth *et al.* 1990; Charlesworth and Hughes 1998; Deng and Lynch 1996). In highly selfing species, recessive lethals will be quickly purged from the population (Lande and Schemske 1985), and will thus contribute little to the different effects of inbreeding versus outbreeding on fitness components. In addition, such populations will have few polymorphic loci with allelic variability maintained by overdominance (Kimura and Ohta 1971; Charlesworth and Charlesworth 1995). Thus, we can reasonably neglect lethal and segregational loads and assume that heterosis in highly inbred populations is due solely to detrimental mutational load. The rate for independently acting detrimental mutations per diploid genome can then be estimated from the formula

$$U_h = -\ln(1 - \delta)/(1 - 2h), \qquad (3)$$

where δ is the reduction in mean fitness of highly selfed individuals compared with randomly mated individuals,

and h is the dominance coefficient associated with a typical detrimental mutation. This method provides underestimates of total mutation rates for reasonable values of h (Charlesworth *et al.* 1990).

The range of values of U_h so obtained from data on heterosis in net fitness for several species of highly selfing plants from three families of angiosperms are shown in the top part of Table 7; where no estimate of h is available, it was assumed to be 0.2 as suggested by the Drosophila data (Crow and Simmons 1983; Charlesworth and Hughes 1998). Because each individual estimate is subject to considerable sampling error, it is probably wise to treat only the median value of approximately 0.3 as meaningful. This estimate is ostensibly independent of the strength of selection against the deleterious mutations, and so may capture a larger fraction of them than the mutation-accumulation method. On the other hand, the fact that h seems to be much closer to 0.5 for mutations with small effects than for mutations with drastic effects (Crow and Simmons 1983) means that very weakly selected mutations are likely to contribute relatively little to heterosis and so would not be detected by this method.

Hughes (1995) and Deng and Lynch (1996, 1997) have suggested an extension of this approach to use the genetic variances of inbred and outbred populations to estimate the decrease in fitness components with inbreeding. By simultaneously estimating the degree of inbreeding depression, U and $\bar{s}$ can also be estimated (Deng and Lynch 1996, 1997). Estimates of U_h applying this method to published data on Drosophila fitness components (B. Charlesworth, unpublished results) and from data on Daphnia (Deng and Lynch 1997) are given in the lower part of Table 7. There is again substantial spread in the individual estimates; the median is approximately 0.34, surprisingly close to the plant value.

While an attractive idea, this method depends heavily on the assumption that mutation-selection balance is the sole force maintaining genetic variation in fitness components. There are good reasons to question this assumption for outbreeding species. In Drosophila, both the genetic variance and inbreeding depression for components of fitness seem to be too high to be explained solely by mutation-selection balance, using the parameter estimates shown in Table 6 (Mukai *et al.* 1974; Charlesworth and Hughes 1998). This implies the existence of a substantial contribution from variation which is actively maintained by selection, causing an upward bias in estimates of U by the Deng-Lynch method. The presence of variation maintained by selection means that U is overestimated to an extent which is difficult to gauge.

The impact of increased rates of mutation: Does the high rate of spontaneous deleterious mutation per sexual generation in multicellular organisms render them sensitive to small rate increases, as seen with RNA vi-

TABLE 7

Indirect estimates of the detrimental mutation rate

Organism	Trait studied	U_h	References
Several species of selfing plants	Various life-history traits	0.12–0.85	Ågren and Schemske (1993), Charlesworth et al. (1994), Johnston and Schoen (1995)
Drosophila melanogaster	Egg-to-adult viability	0.11[a]	Takano et al. (1987)
	Early male mating success	0.33[b]	Hughes (1995)
	Male longevity	0.30[b]	Hughes (1995)
Daphnia pulicaria	Various life-history traits	0.40–0.58	Deng and Lynch (1997)
Daphnia arenata	Various life-history traits	0.19–0.50	Deng and Lynch (1997)

[a] The genomic rate is estimated by assuming that chromosome *2* constitutes 40% of the genome.

[b] The genomic rate is estimated by assuming that chromosome *3* constitutes 48% of the genome.

ruses, retroelements, and mutator strains of *E. coli* and *S. cerevisiae*? We would expect both diploidy, and the infrequent demand for the functions of some genes that mediate responses to environmental challenges, to provide considerable protection from recessive lethal mutations, at least until mutations accumulate to an intolerable frequency. But detrimental mutations seem to have higher levels of dominance than lethals (Crow and Simmons 1983), so that a high genomic mutation rate to detrimentals could imperil an outbred population.

In the 1950s, Wallace (1952, 1956) exposed caged Drosophila populations to continuous radiation accumulating to 250,000 r. The population accumulated a large number of recessive lethals, but its size was not reduced. This is perhaps not surprising in a species with a high reproductive potential. Nevertheless, although heavily mutagenized Drosophila populations showed no overt signs of genetic deterioration, they became weak competitors with nonmutagenized strains (Wijsman 1984). Thus, Wallace's flies were indeed paying a price, but one that would have required a more rigorous environment to reveal.

Recent results with mammals are instructive. Several strains of mice have been rendered homozygous for defects in the mismatch repair genes *Msh2* (De Wind *et al.* 1995; Reitmair *et al.* 1995), *Pms2* (Baker *et al.* 1995) or *Mlh1* (Baker *et al.* 1996; Edelmann *et al.* 1996). These mice display a mutator phenotype in somatic cells, the mutability of microsatellite sequences being increased by roughly 10^2-fold. In the case of *Pms2*, the somatic mutability of an artificial mutational target in a shuttle vector increased about 100-fold throughout the gene (Narayanan *et al.* 1997); most of the mutations were ±1 b, that is, frameshifts that are likely to be null mutations. Most of these mutator mice are superficially healthy, although cancer-prone. Most are sterile, apparently for reasons more mechanical than mutational, but *Pms2*$^{-/-}$ females are fertile. In humans, samples of normal tissue from several cancer patients were defective in *hPMS2* or *hMLH1* (the human homologs of murine *Pms2* and *Mlh1*, respectively; see Parsons *et al.* 1995), suggesting that mutator humans occur naturally and have proper-

ties similar to those of mutator mice. These results suggest, as mentioned above, weaker selection against somatic than against germinal mutations.

Thus, substantially increased somatic mutation rates appear to be compatible with mammalian development. However, a persistently high germline mutation rate would be expected to extinguish the population within a few generations; one can easily imagine mouse breeding schemes that could explore the effects of mutation accumulation under highly mutagenic conditions.

Aging is an important aspect of mammalian development. Aging has long been conjectured to reflect the accumulation of somatic mutations (see Finch 1990). The lack of obviously accelerated aging in mice with 100-fold increased somatic mutation rates considerably weakens this hypothesis.

THE EVOLUTION OF MUTATION RATES

Discernible patterns: Rates of spontaneous mutation display several distinct patterns across taxa. RNA-based organisms have the highest genomic rate per genome replication, $\mu_g \approx 1$–2 for lytic RNA viruses and $\mu_g \approx 0.1$–0.2 for retroviruses and a retrotransposon exclusive of their tenure as parts of host chromosomes. For DNA-based microbes, $\mu_g = 0.0034 = 1/300$. For higher eukaryotes, the mutation rate is properly expressed per *effective* genome, which includes only those parts of the genome in which most mutations produce deleterious effects upon which selection can act effectively; based on rather incomplete information, $\mu_{eg} \approx 0.006$ (range 0.004–0.014), these values presently being indistinguishable from 0.003. The corresponding mutation rate per effective genome per sexual generation (μ_{egs}) varies by at least 40-fold (0.036–1.6), and the mutation rate per total higher eukaryotic genome must considerably exceed 1 in some cases. The lower bound estimates of rates of deleterious mutation per genome per sexual generation (U_h) are about 0.1–0.35 for Arabidopsis and Drosophila, and U_d therefore approaches or somewhat exceeds 1 for these organisms; it may be lower for *C. elegans*. As we have discussed, these estimates need

to be viewed with caution, and may be substantially revised in the future.

Miscellaneous puzzles and rejoinders: Why do the lytic RNA viruses put up with rates of spontaneous mutation so high as to genetically degrade a substantial fraction of their progeny? One speculation has been that these high rates facilitate escape from immune surveillance and other host defenses. This speculation is faulted by the observation that the RNA-based bacteriophage Qβ and its DNA-based counterpart bacteriophage M13, both infecting *E. coli* and displaying similar life histories, nevertheless retain the mutation rate characteristic of their class. Another speculation has been that the replication of RNA viruses *cannot* be more accurate because their replicases lack a proofreading function. However, the retroviruses also lack such a function, but nevertheless achieve a replication accuracy that results in far fewer defective progeny. One general, simple way to increase accuracy would be to decrease the ability of the polymerase to extend from a mismatch, thus aborting mutant progeny. Another, more radical way would be to appropriate a DNA proofreading activity and adapt it to RNA substrates, although this would involve enlarging the genome size and thus increasing the chemical lability of the genome. Yet another speculation is that the benefits from replicating as rapidly as possible outweigh the costs of a high error rate. However, this problem could be solved, as it already has in both prokaryotes and eukaryotes, by employing multiple sites of replication initiation. A final speculation is thus far unfaulted: the retroviral mutation rate is determined primarily by the error rate of transcription, on which the virus cannot improve.

Another characteristic of the RNA viruses, and to a lesser extent the retroviruses, is their extraordinary fecundity. Yields per infected cell of 10^3 to 10^4 iu (and perhaps tenfold more physical particles) are common. Thus, there is a substantial probability that an infected cell will release numerous particles that are free of deleterious mutations. A general property of these viruses that may also bear on their mutation rate is the inherent chemical lability of the RNA backbone, which appears to limit RNA virus genomes to <40,000 b. Thus, larger genomes do not persist and therefore cannot experience more intense selection for reduced μ_b.

Recombination provides a mechanism that may allow a heavily mutagenized population to randomly generate more fit genomes (Pressing and Reanney 1984; Nee 1987). The genome of influenza virus is extensively subdivided and recombines freely, whereas recombination in poliovirus is rare. However, the two viruses have indistinguishable mutation rates, so that recombination rate appears not to be a determinant of mutation rates in this case. Rates of recombination among DNA-based microbes also vary greatly, and thus do not seem to be important in maintaining the strongly invariant μ_g observed in these organisms.

Two deep differences within and across taxa deserve emphasis here. First, selection on mutation rates must operate differently in organisms with rare or no sex (or with full selfing) than in those sexual organisms with frequent outcrossing, because the products of mutation remain coupled to the rate determinants in the former case but are rapidly decoupled in the latter. Second, two quite different mutation rates operate in multicellular organisms: germline rates and somatic rates. We noted that the murine male germline rate appears to be lower than the somatic rate, suggesting that these two rates evolved somewhat differently. In addition, the remarkable lack of immediate phenotypic consequences of mutator mutations in mice and men (except for higher cancer rates) suggests that in the soma, the buffering provided by diploidy protects strongly against the consequences of mutations (Orr 1995).

Evolutionary forces shaping mutation rates: The different patterns of mutation rates among taxa indicate clearly that the rate of mutation is subject to evolutionary change. Because the fidelity of DNA replication depends on elaborate enzymatic machinery, mutational inactivation of any component of which can greatly elevate the mutation rate, selection acts primarily to reduce the standard mutation rate, although allowing higher rates in specific circumstances. Such selection pressure was first posited by Sturtevant (1937), who observed that the vast majority of spontaneous mutations decrease rather than increase fitness. He suggested that the pressure of deleterious mutations would favor selection of genetic modifiers that reduce the mutation rate and thereby reduce the genetic load of deleterious alleles maintained in the population by mutation-selection balance. This led him to ask why the mutation rate does not fall to zero. Some 30 years later, Kimura (1967) suggested that the cost of continually reducing mutation rates would eventually be balanced by what he called the "physiological cost" of doing so.

Selection for modifiers of mutation rates: More generally, we may now inquire what evolutionary factors determine the mutation rate of a species. While we cannot pretend that this question can yet be fully answered, the main ingredients of an answer are now reasonably clear. A well developed body of theory predicts the effect of selection on a modifier gene that causes a small reduction in the genomic mutation rate to deleterious alleles (U_h or U_d in the terminology introduced above), both for freely recombining sexual populations and for completely asexual populations (Kondrashov 1988, 1995). This reduction comes about because a modifier allele causing a reduction in the mutation rate becomes associated in the population with genomes that have a lower-than-average number of deleterious mutations. In the case of a freely recombining diploid species, the selection coefficient on a modifier which reduces the diploid mutation rate by δU is approximately $s\delta U$, where s is the mean selection coefficient against a heterozygous

deleterious mutation. In a diploid asexual species, the selection coefficient is approximately δU. In a completely selfing population, the value is $0.5\,\delta U$, provided that deleterious mutations are not completely recessive. These results assume that the populations in questions are close to deterministic equilibrium under mutation and selection, and so do not take into account the presumably numerous class of deleterious mutations whose effects on fitness are of the order of the reciprocal of the effective population size. To take this class of mutations into account, further studies are needed of the dynamics of selection on mutation rate modifiers.

The intuitive reason for this effect of breeding system is as follows. In an asexual or selfing population, the complete linkage of a mutation-rate modifier to its targets means that the selection pressure on the modifier is determined by the difference between the equilibrium mean fitness of clones containing the modifier allele and of clones carrying its rival allele (Leigh 1970, 1973). The equilibrium mean fitness of a diploid asexual population subject to deleterious mutation at a rate U_d is e^{-Ud} (Haldane 1937; Kimura and Maruyama 1966); thus, the difference in mean fitness between populations carrying the modifier and rival populations is expected to equilibrate at $e^{-(Ud-\delta Ud)} - e^{-Ud} \approx \delta U_d$. With free recombination, on the other hand, a neutral allele remains associated with a mutation that has occurred in the same gamete as itself for an average of only two generations, so that the apparent fitness reduction to the allele due to its initial association with the mutation is $2s$ (Kimura 1967). If a mutation-rate modifier reduces the diploid genomic mutation rate by δU_d, the modifier thus gains an advantage over the rest of the population of approximately $s\delta U_d$ (Kimura 1967; Kondrashov 1995), because only half of the mutations will occur in the same haploid genome as the modifier allele.

It is relatively easy to extend these arguments to include a direct fitness cost to a reduction in the mutation rate. Such a cost prevents the mutation rate from being reduced to zero, for instance because of the energetic costs of diverting cellular resources to proofreading mechanisms (Kimura 1967; Kondrashov 1988, 1995). If such fidelity costs are included in the equations, one can determine the evolutionary equilibrium mutation rate; at this rate, alleles modifying the mutation rate in either direction are neutral or disadvantageous (Kondrashov 1995; McVean and Hurst 1997). While there is experimental evidence for the existence of such costs (Kirkwood *et al.* 1986), there is little direct evidence that U is in fact determined in this way. The paucity of evidence in microbes for mutant alleles with clear-cut antimutator effects suggests that mutation rates are in fact near the physico-chemical minimum that can be achieved at an acceptable cost. The increased resistance to radiation damage in Drosophila laboratory populations that have been exposed to radiation for a long time (Nöthel 1987) is one of the few pieces of evidence that might argue the contrary, but even here the spontaneous mutation rate did not seem to be affected. The existence of Drosophila strains with rates of recessive-lethal mutations that differ by 14-fold (Woodruff *et al.* 1984) shows that not all mutate at the minimum average rate, although the higher Drosophila rates may merely reflect the transitory impact of a transposon infection.

The effect of selection pressure is thus primarily to reduce the genomic mutation rate. The strength of selection on a modifier with a given effect on the mutation rate per base is likely to be proportional to the size of the portion of the genome that produces deleterious mutations, because the same percentage effect of an antimutator gene on the mutation rate per b produces a bigger change in U in larger genomes. δU is thus likely to be proportional to U if the rate of mutation per base is the appropriate scale for measuring effects of mutator genes. If the cost of a given level of reduction in the mutation rate per b is independent of genome size, evolution should result in a rough equality of deleterious mutation rates per genome across species with comparable breeding systems. Such equality is not well supported by the data on RNA viruses in Tables 2 and 3, where the retroviruses have tenfold lower mutation rates per genome generation than the lytic viruses, but is reasonably consistent with the data in Table 4 on DNA-based microbes. The data in Table 5 on higher eukaryotes show fair agreement in the mutation rate per effective genome per replication, but not for the rate per generation as predicted theoretically, with *C. elegans* being especially diverged from the others (see also Table 6).

A possible explanation for different U values in different taxa is that the cost of fidelity may vary with life history or genome size. The total energetic cost of a given change in fidelity per base per generation is likely be greater in species with larger effective genome sizes or more germ-line cell divisions per generation, so that a complete equality of deleterious mutation rates per genome is unlikely to be achieved. The fact that humans appear to have the highest rate of mutation per effective genome per generation (μ_{egs} in Table 5) could be explained by either or both of these effects, among other possibilities.

Sexual versus asexual species: There is clearly much stronger selection in favor of reducing the mutation rate in asexual or selfing organisms than in sexual species. Data from Drosophila suggest that s is of the order of 0.01–0.02 (Crow and Simmons 1983; Charlesworth and Hughes 1998), so that that the difference can easily be as high as two orders of magnitude. If the only factors controlling the fate of modifiers that reduce mutation rates are the fitness advantage of a reduction in mutational load and the cost of increased fidelity of replication, then one might expect to see a much higher mutation rate per genome in sexual compared with asexual

or self-fertilizing taxa, after correcting for any differences in effective genome size. But a variety of factors lead asexual or selfing populations to have short evolutionary persistence times (Kondrashov 1993), so that there may not be enough time for them to evolve lower mutation rates. The low mutation rate per effective genome per generation in *C. elegans* (Table 5) may reflect the high degree of self-fertilization in this species. The fact that effective mutation rates in different species of bacteria and lower eukaryotes are similar (Table 4), despite wide variations in genome size and mode of reproduction, suggest either that differences in breeding system do not in fact matter very much, or that there is enough recombination in ostensibly asexual organisms such as bacteria (Maynard Smith 1991) that the differences in breeding system are more apparent than real. If the latter is the case, then the much higher effective mutation rates in higher eukaryotes (Table 5) can only be explained in the present model by a greatly increased cost of fidelity of replication per generation. Unfortunately, quantitative data on this cost are completely lacking, so that this conclusion remains speculative.

In an asexual species, deleterious genes are eliminated from the population in the same genotypes in which they occur. If mutations occur independently, they are eliminated independently. In a sexual species, deleterious mutations are regrouped every generation, so that it is possible in principle for mutations to be eliminated in groups. Is there a Maxwell's Demon who ensures that each "genetic death" picks off several mutations? Directional epistasis or quasi-truncation selection can have such an effect (Crow 1997), but how effective this is in natural populations is an open question. It is possible that a mutation rate that would lead to extinction in an asexual species would be tolerated in a sexual one.

Adaptive mutations: It has often been suggested that higher mutation rates would be adaptive in populations undergoing strong directional selection, because mutational variability would enhance the speed of response to such a selection pressure. Sturtevant in fact raised this possibility in 1937, only to dismiss it with the phrase, "While this effect may occur, it is difficult to imagine its operation." Undeterred, Kimura (1960, 1967) developed a theory of selection on the mutation rate, according to which the genetic load experienced by a population from the joint effects of deleterious mutations and the substitution of alleles due to selection in a changing environment is assumed to be minimized by selection on the mutation rate. He showed that the optimal effective genomic mutation rate according to this criterion is equal to the rate of substitution of favorable alleles in the genome as a whole (Kimura 1967). The weakness of this theory for sexually reproducing organisms is that it implicitly assumes the operation of group rather than individual selection, with all the attendant difficulties (Leigh 1970, 1973).

The analysis of models of selection on genes that

modify the mutation rate has allowed progress on this problem. As we have seen for the case of mutations with purely deleterious effects, recombination in a sexual species greatly weakens the force of selection on a mutation-rate modifier. This problem is more acute for the case of favorable mutations, because these are likely to be much rarer than deleterious mutations. A modifier allele that increases the mutation rate may thus receive a short-lived boost in frequency from its association with a favored allele that it has induced, but soon loses this advantage as a result of recombination (Kimura 1967; Leigh 1970, 1973). The pressure of selection for an increased mutation rate from the induction of favorable mutations is thus extremely weak in a sexually reproducing organism, and is likely to be overwhelmed by the disadvantage of deleterious mutations.

The situation is somewhat different in species with low levels of recombination, such as many bacteria, because a modifier can remain associated with a mutation that it has induced. In the absence of recombination, Kimura's results apply to the process of selection on favorable mutations (Leigh 1970, 1973). Populations that have experienced a severe challenge from a novel environment might therefore be expected to show an increased frequency of mutator alleles (Taddei *et al.* 1997). Several experiments with bacterial populations support this conclusion, and the genetic basis of the increased mutation rate has been identified in some cases (Sniegowski *et al.* 1997). This is consistent with the observation cited earlier that some natural isolates of bacteria harbor mutator genes. A mutator allele still faces a long-term problem because it causes a higher load of deleterious alleles, so that once adaptation to the new environment has occurred, selection for a reduction in the mutation rate will be renewed. A mutator strain of bacteria has been observed to evolve a lower mutation rate when grown in a chemostat for over 2,000 generations (Tröbner and Piechocki 1984), presumably as a result of selection of this kind.

One special circumstance in which a high mutation rate is favored is when an organism is confronted with a rapidly cycling or otherwise continually changing environment, so that it pays to be able to respond by producing novel genotypes at loci which are responsible for adaptation to the new state of the environment (Gillespie 1981a; Ishii *et al.* 1989). Polymorphisms maintained by selection in a constant environment can also induce selection for increased mutation rates (Gillespie 1981b; Kondrashov 1995). If increased mutational load is to be avoided, then hypermutability should be targeted to specific loci or should be transient. The responses of several pathogens to the host's immune system by antigen switching are an excellent example of this, although they are achieved by special genetic devices rather than conventional mutagenesis (Sasaki 1994).

On the mechanisms of mutation and mutation prevention: Organisms limit their mutation rates by diverse

mechanisms. These include metabolic controls over concentrations of endogenous and exogenous mutagens, pre-replication DNA repair systems, the insertion accuracy of polymerases, 3′-exonucleolytic proofreading, and several post-replication systems for repairing mismatches. Different organisms apply different sets of these mechanisms, and the efficiency of a particular mechanism varies among organisms. Sometimes an organism's mutation rate is considered to be "determined" by the particular set of mechanisms it applies. It is more accurate, however, to view that organism's mutation rate as "determined" by deep evolutionary forces, by the life history it has adopted, and by accidents of its evolutionary history. The particular mechanisms employed and their efficiencies are merely devices to carry out the underlying necessity.

Note added in proof: Howe and Ares (1997) report that yeast introns employ stem-loop structures to bring together the splice junctions at the opposite ends of an intron and thus to increase the probability of correct splicing. If this were a general effect, it might add 10^2 to 10^3 b to the mutational target size of a locus. Such an effect would have little or no effect on the values of μ_{eg} or μ_{egs} in Table 5, but would slightly increase the values of G_e and slightly reduce the values of μ_b and μ_g. These changes would be small relative to the present uncertainties of these values in higher eukaryotes.

LITERATURE CITED

Ågren, J., and D. W. Schemske, 1993 Outcrossing rate and inbreeding depression in two annual monoecious herbs, *Begonia hirsuta* and *B. semiovata.* Evolution **47:** 125–135.

Akiyama, M., S. Kyoizumi, Y. Hirai, Y. Kusunoki, K. Iwamoto *et al.*, 1995 Mutation frequency in human blood cells increases with age. Mutat. Res. **338:** 141–149.

Ashburner, M., 1989 *Drosophila, A Laboratory Handbook.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Auerbach, C., 1959 Spontaneous mutations in dry spores of *Neurospora crassa.* Z. Vererbungsl **90:** 335–346.

Baker, S. M., C. E. Bronner, L. Zhang, A. W. Plug, M. Robatzek *et al.*, 1995 Male mice defective in the DNA mismatch repair gene *PMS2* exhibit abnormal chromosome synapsis in meiosis. Cell **82:** 309–319.

Baker, S. M., A. W. Plug, T. A. Prolla, C. E. Bronner, A. C. Harris *et al.*, 1996 Involvement of mouse *Mlh1* in DNA mismatch repair and meiotic crossing over. Nature Genet. **13:** 336–342.

Bateman, A. J., 1959 The viability of near-normal irradiated chromosomes. Int. J. Radiat. Biol. **1:** 170–180.

Benian, G. M., J. E. Kiff, N. Neckelmann, D. G. Moerman and R. H. Waterston, 1989 Sequence of an unusually large protein implicated in regulation of myosin activity in *C. elegans.* Nature **342:** 45–50.

Benzer, S., 1961 On the topography of the genetic fine structure. Proc. Natl. Acad. Sci. USA **47:** 403–415.

Bird, A. P., 1995 Gene number, noise reduction and biological complexity. Trends Genet. **11:** 94–100.

Blattner, F. R. III, G. Plunkett, C. A. Block, N. T. Perna, V. Burland *et al.*, 1997 The complete genome sequence of *Escherichia coli* K-12. Science **277:** 1453–1474.

Chao, L., and E. C. Cox, 1983 Competition between high and low mutating strains of *Escherichia coli.* Evolution **37:** 125–134.

Charlesworth, B., and K. A. Hughes, 1998 The maintenance of genetic variation in life-history traits, in *Evolutionary Genetics from Molecules to Morphology,* edited by R. S. Singh and C. B. Krimbas. Cambridge University Press, Cambridge (in press).

Charlesworth, B., D. Charlesworth and M. T. Morgan, 1990 Genetic loads and estimates of mutation rates in very inbred plant populations. Nature **347:** 380–382.

Charlesworth, D., and B. Charlesworth, 1995 Quantitative genetics in plants: the effect of breeding system on genetic variability. Evolution **49:** 911–920.

Charlesworth, D., E. E. Lyons and L. B. Litchfield, 1994 Inbreeding depression in two highly inbreeding populations of *Leavenworthia.* Proc. R. Soc. Lond. Ser. B Biol. Sci. **258:** 209–214.

Christensen, R. B., J. R. Christensen and C. W. Lawrence, 1985 Conjugation-dependent enhancement of induced and spontaneous mutation in the *lacI* gene of *E. coli.* Mol. Gen. Genet. **201:** 35–37.

Clark, D. V., T. M. Rogalski, L. M. Donati and D. L. Baillie, 1988 The *unc-22* (*IV*) region of *Caenorhabditis elegans.* Genetics **119:** 345–353.

Connor, K. F., and R. M. Lanner, 1991 Effects of tree age on pollen, seed, and seedling characteristics in Great Basin bristlecone pine. Bot. Gaz. **152:** 114–122.

Crow, J. F., 1993 How much do we know about spontaneous human mutation rates? Environ. Mol. Mutagen. **21:** 122–129.

Crow, J. F., 1995 Spontaneous mutation as a risk factor. Exp. Clin. Immunogenet. **12:** 121–128.

Crow, J. F., 1997 The high spontaneous mutation rate: is it a health risk? Proc. Natl. Acad. Sci. USA **94:** 8380–8386.

Crow, J. F., and M. J. Simmons, 1983 The mutation load in Drosophila., pp. 1–35 in *The Genetics and Biology of Drosophila, Vol. 3c,* edited by M. Ashburner, H. L. Carson and J. N. Thompson. Academic Press, London.

Crumpacker, D. W., 1967 Genetic loads in maize (*Zea mays* L.) and other cross-fertilized plants and animals. Evol. Biol. **1:** 306–423.

Deng, H.-W., and M. Lynch, 1996 Estimation of deleterious-mutation parameters in natural populations. Genetics **144:** 349–360.

Deng, H.-W., and M. Lynch, 1997 Inbreeding depression and inferred deleterious-mutation parameters in Daphnia. Genetics **147:** 147–155.

de Wind, N., M. Dekker, A. Berns, M. Radman and H. te Riele, 1995 Inactivation of the mouse *Msh2* gene results in mismatch repair deficiency, methylation tolerance, hyperrecombination, and predisposition to cancer. Cell **82:** 321–330.

Drake, J. W., 1966 Spontaneous mutations accumulating in bacteriophage T4 in the complete absence of DNA replication. Proc. Natl. Acad. Sci. USA **55:** 738–743.

Drake, J. W., 1991 A constant rate of spontaneous mutation in DNA-based microbes. Proc. Natl. Acad. Sci. USA **88:** 7160–7164.

Drake, J. W., 1993a Rates of spontaneous mutation among RNA viruses. Proc. Natl. Acad. Sci. USA **90:** 4171–4175.

Drake, J. W., 1993b General antimutators are improbable. J. Mol. Biol. **229:** 8–13.

Drake, J. W., and L. S. Ripley, 1994 Mutagenesis, pp. 98–124 in *Molecular Biology of Bacteriophage T4,* edited by J. D. Karam. American Society for Microbiology, Washington, D.C.

Drost, J. B., and W. R. Lee, 1995 Biological basis of germline mutation: comparisons of spontaneous germline mutation rates among Drosophila, mouse, and human. Environ. Mol. Mutagen. **25,** Suppl. **26:** 48–64.

Edelmann, W., P. E. Cohen, M. Kane, K. Lau, B. Morrow *et al.*, 1996 Meiotic pachytene arrest in MLH1-deficient mice. Cell **85:** 1125–1134.

Eide, D., and P. Anderson, 1985 The gene structures of spontaneous mutations affecting a *Caenorhabditis elegans* myosin heavy chain gene. Genetics **109:** 67–79.

Fernandez, J., and C. Lopez-Fanjul, 1996 Spontaneous mutational variances and covariances for fitness-related traits in *Drosophila melanogaster.* Genetics **143:** 829–837.

Finch, C. E., 1990 *Longevity, Senescence, and the Genome.* University of Chicago Press, Chicago, IL.

Finnegan, D. J., and D. H. Fawcett, 1986 Transposable elements in *Drosophila melanogaster.* Oxf. Surv. Eukaryotic Genes **3:** 1–62.

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

Foster, P. L., 1997 Nonadaptive mutations occur on the F′ episome

during adaptive mutation conditions in *Escherichia coli.* J. Bacteriol. **179:** 1550–1554.

Gabriel, A., M. Willems, E. H. Mules and J. D. Boeke, 1996 Replication infidelity during a single cycle of Ty*1* retrotransposition. Proc. Natl. Acad. Sci. USA **93:** 7767–7771.

Garcia-Dorado, A., 1997 The rate and effects distribution of viability mutations in Drosophila: minimum distance estimation. Evolution **51:** 1130–1139.

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

Gillespie, J. H., 1981b Evolution of the mutation rate at a heterotic locus. Proc. Natl. Acad. Sci. USA **78:** 2452–2454.

Goelet, P., G. P. Lomonossoff, P. J. G. Butler, M. E. Akam, M. J. Gait *et al.*, 1982 Nucleotide sequence of tobacco mosaic virus RNA. Proc. Natl. Acad. Sci. USA **79:** 5818–5822.

Goffeau, A., B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon *et al.*, 1996 Life with 6000 genes. Science **274:** 546–567.

Greenwald, I. S., and H. R. Horvitz, 1980 *unc-93(e1500)*: a behavioral mutant of *Caenorhabditis elegans* that defines a gene with a wild-type null phenotype. Genetics **96:** 147–164.

Gross, M. D., and E. C. Siegel, 1981 Incidence of mutator strains in *Escherichia coli* and coliforms in nature. Mutat. Res. **91:** 107–110.

Hall, J. D., D. M. Coen, B. L. Fisher, M. Weisslitz, S. Randall *et al.*, 1984 Generation of genetic diversity in herpes simplex virus: an antimutator phenotype maps to the DNA polymerase locus. Virology **132:** 26–37.

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

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

Holland, J. J., E. Domingo, J. C. de la Torre and D. A. Steinhauer, 1990 Mutation frequencies at defined single codon sites in vesicular stomatitis virus and poliovirus can be increased only slightly by chemical mutagenesis. J. Virol. **64:** 3960–3962.

Houle, D., 1992 Comparing evolvability and variability of quantitative traits. Genetics **130:** 195–204.

Howe, K. J., and M. Ares, Jr, 1997 Intron self-complementarity enforces exon inclusion in yeast pre-mRNA. Proc. Natl. Acad. Sci. **94:** 12467–12472.

Hughes, K. A., 1995 The inbreeding decline and average dominance of genes affecting male life-history characters in *Drosophila melanogaster*. Genet. Res. **65:** 41–52.

Ishii, K., H. Matsuda, Y. Iwasa and A. Sasaki, 1989 Evolutionarily stable mutation rate in a periodically changing environment. Genetics **121:** 163–174.

Jacobs, K. L., and D. W. Grogan, 1997 Rates of spontaneous mutation in an Archaeon from geothermal environments. J. Bacteriol. **179:** 3298–3303.

Johnston, M. O., and D. J. Schoen, 1995 Mutation rates and dominance levels of genes affecting total fitness in two angiosperm species. Science **267:** 226–229.

Jyssum, K., 1960 Observations on two types of genetic instability in *Escherichia coli.* Acta Path. Microbiol. Scand. **48:** 113–120.

Karn, J., S. Brenner and L. Barnett, 1983 Protein structural domains in the *Caenorhabditis elegans unc-54* myosin heavy chain gene are not separated by introns. Proc. Natl. Acad. Sci. USA **80:** 4253–4257.

Kearney, C. M., J. Donson, G. E. Jones and W. O. Dawson, 1993 Low level of genetic drift in foreign sequences replicating in a RNA virus in plants. Virology **192:** 11–17.

Keightley, P. D., 1994 The distribution of mutation effects on viability in *Drosophila melanogaster*. Genetics **138:** 1–8.

Keightley, P. D., 1996 The nature of deleterious mutation load in Drosophila. Genetics **144:** 1993–1999.

Keightley, P. D., and A. Caballero, 1997 Genomic mutation rate for fitness in *Caenorhabditis elegans.* Proc. Natl. Acad. Sci. USA **94:** 3823–3827.

Khush, G. S., and C. M. Rick, 1967 Studies on the linkage map of chromosome 4 of the tomato and on the transmission of induced deficiencies. Genetica **38:** 74–94.

Kibota, T. T., and M. Lynch, 1996 Estimate of the genomic mutation rate deleterious to overall fitness in *E. coli.* Nature **381:** 694–696.

Kimble, J., and S. Ward, 1988 Germ-line development and fertilization, pp. 191–213 in *The Nematode Caenorhabditis elegans*, edited by W. B. Wood. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Kimura, M., 1960 Optimum mutation rate and degree of dominance as determined by the principle of minimum genetic load. J. Genet. **57:** 21–34.

Kimura, M., 1967 On the evolutionary adjustment of spontaneous mutation rates. Genet. Res. **9:** 23–34.

Kimura, M., 1983a *The Neutral Theory of Molecular Evolution.* Cambridge Univ. Press, Cambridge.

Kimura, M., 1983b Rare variant alleles in the light of the neutral theory. Mol. Biol. Evol. **1:** 84–93.

Kimura, M., and T. Maruyama, 1966 The mutation load with epistatic gene interactions in fitness. Genetics **54:** 1337–1351.

Kimura, M., and T. Ohta, 1971 *Theoretical Topics in Population Genetics.* Princeton University Press, Princeton, N.J.

Kirkwood, T. B. L., R. F. Rosenberger and D. J. Galas, 1986 *Accuracy in Molecular Processes.* Chapman and Hall, London.

Klekowski, E. J., 1973 Genetic load in *Osmunda regalis* populations. Am. J. Bot. **60:** 146–154.

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

Klekowski, E. J., 1992 Mutation rates in diploid annuals—are they immutable? Int. J. Plant Sci. **153:** 462–465.

Klekowski, E. J., and P. J. Godfrey, 1989 Aging and mutation in plants: a comparison of woody mangroves and herbaceous annuals. Nature **340:** 389–391.

Klekowski, E. J., R. L. Lowenfeld and P. K. Hepler, 1994 Mangrove genetics. II. Outcrossing and lower spontaneous mutation rates in Puerto Rican Rhizophora. Int. J. Plant Sci. **155:** 373–381.

Knight, G. R., and A. Robertson, 1957 Fitness as a measurable character in Drosophila. Genetics **42:** 524–530.

Kohler, S. W., G. S. Provost, A. Fieck, P. L. Kretz, W. O. Bullock *et al.*, 1991 Spectra of spontaneous and mutagen-induced mutations in the *lacI* gene in transgenic mice. Proc. Natl. Acad. Sci. USA **88:** 7958–7962.

Kondrashov, A., 1984 Deleterious mutations as an evolutionary factor. I. The advantage of recombination. Genet. Res. **44:** 325–332.

Kondrashov, A. S., 1988 Deleterious mutations and the evolution of sexual reproduction. Nature **336:** 435–440.

Kondrashov, A. S., 1993 A classification of hypotheses on the advantage of amphimixis. J. Hered. **84:** 372–387.

Kondrashov, A. S., 1995 Modifiers of mutation-selection balance: general approach and the evolution of mutation rates. Genet. Res. **66:** 53–69.

Kondrashov, A. S., and J. F. Crow, 1993 A molecular approach to estimating the human deleterious mutation rate. Hum. Mutat. **2:** 229–234.

Kondrashov, A. S., and D. Houle, 1994 Genotype-environment interaction and the estimation of the genomic mutation rate in *Drosophila melanogaster.* Proc. R. Soc. Lond. Ser. B Biol. Sci. **258:** 221–227.

Kunz, B. A., and B. W. Glickman, 1983 The infidelity of conjugal DNA transfer in *Escherichia coli.* Genetics **105:** 489–500.

Kutter, E., T. Stidham, B. Guttman, E. Kutter, D. Batts *et al.*, 1994 Genomic map of bacteriophage T4, pp. 491–518 in *Molecular Biology of Bacteriophage T4*, edited by J. D. Karam. American Society for Microbiology, Washington D.C.

Laird, C.D., 1971 Chromatid structure: relationship between DNA content and nucleotide sequence diversity. Chromosoma **32:** 378–406.

Lande, R., and D. W. Schemske, 1985 The evolution of self fertilization and inbreeding depression in plants. I. Genetic models. Evolution **39:** 24–40.

LeClerc, J. E., B. Li, W. L. Payne and T. A. Cebula, 1996 High mutation frequencies among *Escherichia coli* and *Salmonella* pathogens. Science **274:** 1208–1211.

Leigh, E. G., 1970 Natural selection and mutability. Am. Nat. **104:** 301–305.

Leigh, E. G., 1973 The evolution of mutation rates. Genetics Suppl. **73:** 1–18.

Levin, J. Z., and H. R. Horvitz, 1992 The *Caenorhabditis elegans unc-93* gene encodes a putative transmembrane protein that regulates muscle contraction. J. Cell Biol. **117:** 143–155.

Liu, J., B. Schrank and R. H. Waterston, 1996 Interaction between

a putative mechanosensory membrane channel and a collagen. Science **273:** 361–364.

Mansky, L. M., and H. M. Temin, 1994 Lower mutation rate of bovine leukemia virus relative to spleen necrosis virus. J. Virol. **68:** 494–499.

Mansky, L. M., and H. M. Temin, 1995 Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. J. Virol. **69:** 5087–5094.

Mao, E. F., L. Lane, J. Lee and J. H. Miller, 1997 Proliferation of mutators in a cell population. J. Bacteriol. **179:** 417–422.

Matic, I., M. Radman, F. Taddei, B. Picard, C. Doit et al., 1997 Highly variable mutation rates in commensal and pathogenic Escherichia coli. Science **277:** 1833-1834.

Maynard Smith, J., 1991 The population genetics of bacteria. Proc. R. Soc. Lond. Ser. B Biol. Sci. **245:** 37–41.

McGeoch, D. J., M. A. Dalrymple, A. J. Davison, A. Dolan, M. C. Frame et al., 1988 The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. J. Gen. Virol. **69:** 1531–1574.

McKnight, S. L., 1980 The nucleotide sequence and transcript map of the herpes simplex virus thymidine kinase gene. Nucleic Acids Res. **8:** 5949–5964.

McVean, G. T., and L. D. Hurst, 1997 Evidence for a selectively favorable reduction in the mutation rate of the X chromosome. Nature **386:** 388–392.

Moxon, E. R., P. B. Rainey, M. A. Nowak and R. E Lenski, 1994 Adaptive evolution of highly mutable loci in pathogenic bacteria. Curr. Biol. **4:** 23–33.

Mukai, T., 1964 The genetic structure of natural populations of Drosophila melanogaster. I. Spontaneous mutation rate of polygenes controlling viability. Genetics **50:** 1–19.

Mukai, T., 1969 The genetic structure of natural populations of Drosophila melanogaster. VII. Synergistic interaction of spontaneous mutant polygenes controlling viability. Genetics **61:** 749–761.

Mukai, T., and C. C. Cockerham, 1977 Spontaneous mutation rates at enzyme loci in Drosophila melanogaster. Proc. Natl. Acad. Sci. USA **74:** 2514–2517.

Mukai, T., S. I. Chigusa, L. E. Mettler and J. F. Crow, 1972 Mutation rate and dominance of genes affecting viability in Drosophila melanogaster. Genetics **72:** 335–355.

Mukai, T., R. A. Cardellino, T. K. Watanabe and J. F. Crow, 1974 The genetic variance for viability and its components in a local population of Drosophila melanogaster. Genetics **78:** 1195–1206.

Muller, H. J., 1928 The measurement of gene mutation rate in Drosophila, its high variability, and its dependence on temperature. Genetics **13:** 279–357.

Muller, H. J., 1952 The standard error of the frequency of mutants some of which are of common origin. Genetics **37:** 608.

Narayanan, L., J. A. Fritzell, S. M. Baker, R. M. Liskay and P. M. Glazer, 1997 Elevated levels of mutation in multiple tissues of mice deficient in the DNA mismatch repair gene *Pms2.* Proc. Natl. Acad. Sci. USA **94:** 3122–3127.

Nee, S., 1987 The evolution of multicompartmental genomes in viruses. J. Mol. Evol. **25:** 277–281.

Ninio, J., 1991 Transient mutators: a semiquantitative analysis of the influence of translation and transcription errors on mutation rates. Genetics **129:** 957–962.

Nöthel, H., 1987 Adaptation of *Drosophila melanogaster* populations to high mutation pressure: evolutionary adjustment of mutation rates. Proc. Natl. Acad. Sci. USA **84:** 1045–1048.

Ohnishi, O., 1977 Spontaneous and ethyl methanesulfonate-induced mutations controlling viability in *Drosophila melanogaster.* II. Homozygous effects of polygenic mutations. Genetics **87:** 529–545.

Ohnishi, O., 1982 Population genetics of cultivated buckwheat, *Fagopyrum esculentum* Moench. I. Frequency of chlorophyll-deficient mutants in Japanese populations. Jpn. J. Genet. **57:** 623–639.

Ohta, T., and C. C. Cockerham, 1974 Detrimental genes with partial selfing and effects on a neutral locus. Genet. Res. **23:** 191–200.

Orr, H. A., 1995 Somatic mutation favors the evolution of diploidy. Genetics **139:** 1441–1447.

Parsons, R., G.-M. Li, M. Longley, P. Modrich, B. Liu et al., 1995 Mismatch repair deficiency in phenotypically normal human cells. Science **268:** 738–740.

Parthasarathi, S., A. Varela-Echavarría, Y. Ron, B. D. Preston

and J. P. Dougherty, 1995 Genetic rearrangements occurring during a single cycle of murine leukemia virus vector replication: characterization and implications. J. Virol. **69:** 7991–8000.

Pathak, V. K., and H. M. Temin, 1990 Broad spectrum of *in vivo* forward mutations, hypermutations, and mutational hotspots in a retroviral shuttle vector after a single replication cycle: substitutions, frameshifts, and hypermutations. Proc. Natl. Acad. Sci. USA **87:** 6019–6023.

Peterson, S. N., T. Lucier, K. Heitzman, E. A. Smith, K. F. Bott et al., 1995 Genetic map of the Mycoplasma genitalium chromosome. J. Bacteriol. **177:** 3199–3204.

Pressing, J., and D. C. Reanney, 1984 Divided genomes and intrinsic noise. J. Mol. Evol. **20:** 135–146.

Quiñones, A., and R. Piechocki, 1985 Isolation and characterization of *Escherichia coli* antimutators. Mol. Gen. Genet. **201:** 315–322.

Ratner, L., W. Haseltine, R. Patarca, K. J. Livak, B. Starcich et al., 1985 Complete nucleotide sequence of the AIDS virus, HTLV-III. Nature **313:** 277–284.

Reha-Krantz, L. J., R. L. Nonay and S. Stocki, 1993 Bacteriophage T4 DNA polymerase mutations that confer sensitivity to the PPi analog phosphonoacetic acid. J. Virol. **67:** 60–66.

Reitmair, A. H., R. Schmits, A. Ewel, B. Bapat, M. Redston et al., 1995 MSH2 deficient mice are viable and susceptible to lymphoid tumours. Nature Genet. **11:** 64–70.

Risch, N, E. W. Reich, M. M. Wishnick and J. G. McCarthy, 1987 Spontaneous mutation and parental age in humans. Am. J. Hum. Genet. **41:** 218–248.

Rosenbluth, R. E., C. Cuddeford and D. L. Baillie, 1983 Mutagenesis in *Caenorhabditis elegans.* I. A rapid eukaryotic mutagen test system using the reciprocal translocation *eT1(III; V).* Mutat. Res. **119:** 39–48.

Russell, L. B., and W. L. Russell, 1996 Spontaneous mutations recovered as mosaics in the mouse specific-locus test. Proc. Natl. Acad. Sci. USA **93:** 13072–13077.

Sagata, N., T. Yasunaga, J. Tsuzuku-Kawamura, K. Ohishi, Y. Ogawa et al., 1985 Complete nucleotide sequence of the genome of bovine leukemia virus: its evolutionary relationship to other retroviruses. Proc. Natl. Acad. Sci. USA **82:** 677–681.

Sapienza, C., 1994 Parental origin effects, genome imprinting, and sex-ratio distortion: double or nothing? Am. J. Hum. Genet. **55:** 1073–1075.

Sasaki, A., 1994 Evolution of antigen drift/switching: continuously evading pathogens. J. Theor. Biol. **168:** 291–308.

Schalet, A., 1960 A study of spontaneous visible mutations in *Drosophila melanogaster.* Ph.D. Thesis, Indiana University, Bloomington.

Schnell, M. J., L. Buonocore, M. A. Whitt and J. K. Rose, 1996 The minimal conserved transcription stop-start signal promotes stable expression of a foreign gene in vesicular stomatitis virus. J. Virol. **70:** 2318–2323.

Shiang, R., L. M. Thompson, Y.-Z. Zhu, D. M. Church. T. J. Fielder et al., 1994 Mutations in the transmembrane domain of FGFR3 cause the most common genetic form of dwarfism, achondroplasia. Cell **78:** 335–342.

Simmons, M. J., E. W. Sheldon and J. F. Crow, 1978 Heterozygous effects on fitness of EMS-treated chromosomes in *Drosophila melanogaster.* Genetics **88:** 575–590.

Sniegowski, P. D., P. J. Gerrish and R. E. Lenski, 1997 Evolution of high mutation rates in experimental populations of *E. coli.* Nature **387:** 703–705.

Stadler, L. J., 1929 Chromosome number and the mutation rate in *Avena* and *Triticum.* Proc. Natl. Acad. Sci. USA **15:** 876–881.

Stadler, L. J., 1930 The frequency of mutation of specific genes in maize (Abstr.). Anat. Rec. **47:** 381.

Steponkus, P. L., and S. Caldwell, 1993 An optimized procedure for the cryopreservation of *Drosophila melanogaster* embryos. Cryo Lett. **14:** 375–380.

Sturtevant, A. H., 1937 Essays on evolution. I. On the effects of selection on the mutation rate. Q. Rev. Biol. **12:** 464–476.

Sulston, J. E., and S. Brenner, 1974 The DNA of *Caenorhabditis elegans.* Genetics **77:** 95–104.

Sved, J. A., 1975 Fitness of third chromosome homozygotes in *Drosophila melanogaster.* Genet. Res. **25:** 197–200.

Taddei, F., M. Radman and J. A. Halliday, 1995 Mutation rate of the F episome. Science **269:** 288–289.

Taddei, F., M. Radman, J. Maynard-Smith, B. Toupance, P. H.

Gouyon *et al.*, 1997 Role of mutator alleles in adaptive evolution. Nature **387:** 700–702.

Takano, T., S. Kusakabe and T. Mukai, 1987 The genetic structure of natural populations of *Drosophila melanogaster*. XX. Comparisons of genotype-environment interaction in viability between a northern and southern population. Genetics **117:** 245–254.

Torkelson, J., R. S. Harris, M.-J. Lombardo, J. Nagendran, C. Thulin *et al.*, 1997 Genome-wide hypermutation in a subpopulation of stationary-phase cells underlies recombination-dependent adaptive mutation. EMBO J. **16:** 3303–3311.

Tröbner, W., and R. Piechocki, 1984 Selection against hypermutability in *Escherichia coli* during long term evolution. Mol. Gen. Genet. **198:** 177–178.

Vogel, F., and A. Motulsky, 1997 *Human Genetics: Problems and Approaches*, Ed. 3. Springer-Verlag, Berlin, New York.

Walker, G. C., 1984 Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*. Microbiol. Rev. **48:** 60–93.

Wallace, B., 1952 The estimation of adaptive values of experimental populations. Evolution **6:** 333–341.

Wallace, B., 1956 Studies on irradiated populations of *Drosophila melanogaster*. J. Genet. **54:** 280–293.

Wijsman, E., 1984 The effect of mutagenesis on competitive ability in Drosophila. Evolution **38:** 571–581.

Willetts, N., and R. Skurray, 1987 Structure and function of the F factor and mechanism of conjugation, pp. 1110–1133 in *Escherichia coli and Salmonella typhimurium*, edited by F. C. Neidhardt. American Society for Microbiology, Washington, D.C.

Woodruff, R. C., J. N. Thompson, M. A. Seeger and W. E. Spivey, 1984 Variation in spontaneous mutation rates in natural population lines of *Drosophila melanogaster*. Heredity **53:** 223–234.

Woodruff, R. C., H. Huai and J. N. Thompson, 1997 Clusters of new mutation in the evolutionary landscape. Genetica **98:** 149–160.

Yamaguchi, O., and T. Mukai, 1974 Variation of spontaneous occurrence rates of chromosomal aberrations in the second chromosomes of *Drosophila melanogaster*. Genetics **78:** 1209–1221.