J. theor. Biol. (1995) **175,** 583–594

Contamination of the Genome by Very Slightly Deleterious Mutations: Why Have We Not Died 100 Times Over?

ALEXEY S. KONDRASHOV

Section of Ecology and Systematics, Cornell University, Ithaca, NY 14853, *U.S.A.*

(*Received on* 21 *October* 1994, *Accepted in revised form on* 3 *February* 1995)

It is well known that when s, the selection coefficient against a deleterious mutation, is below $\sim 1/4N_e$, where N_e is the effective population size, the expected frequency of this mutation is ~ 0.5, if forward and backward mutation rates are similar. Thus, if the genome size, G, in nucleotides substantially exceeds the N_e of the whole species, there is a dangerous range of selection coefficients, $1/G < s < 1/4N_e$. Mutations with s within this range are neutral enough to accumulate almost freely, but are still deleterious enough to make an impact at the level of the whole genome. In many vertebrates $N_e \approx 10^4$, while $G \approx 10^9$, so that the dangerous range includes more than four orders of magnitude. If substitutions at 10% of all nucleotide sites have selection coefficients within this range with the mean 10^{-6}, an average individual carries ~ 100 lethal equivalents. Some data suggest that a substantial fraction of nucleotides typical to a species may, indeed, be suboptimal. When selection acts on different mutations independently, this implies too high a mutation load. This paradox cannot be resolved by invoking beneficial mutations or environmental fluctuations. Several possible resolutions are considered, including soft selection and synergistic epistasis among very slightly deleterious mutations.

© 1995 Academic Press Limited

Introduction

The effects of individual deleterious mutations on fitness vary from lethality to neutrality. If the coefficient of selection, s, against a mutation is substantial, its dynamics are essentially deterministic and depend only on mutation and selection (Wright, 1929; Haldane, 1937; Kimura & Maruyama, 1966). In contrast, dynamics of very slightly deleterious mutations (VSDMs) are stochastic, being also influenced by random drift (see Ohta, 1992). Such alleles can become frequent and even be temporarily fixed, until the best allele is recreated by mutation. This can increase the mutation load, compared to the case of substantially deleterious alleles that are always kept rare by selection (Kimura *et al*., 1963). The study of VSDMs constitutes one of the pillars of population genetics (Crow, 1970; Crow & Kimura, 1970; Kimura, 1983; Ohta, 1992).

E-mail: ask3@crux2.cit.cornell.edu.

A genome can carry a "tremendous" (Li, 1987) number of VSDMs and the consequences of this contamination in the context of molecular evolution have been carefully studied (Ohta, 1992). In contrast, its implications to the fitness of individuals have not attracted much attention. Kimura *et al*. (1963: 1308) suggested that these implications are serious only in rather small populations, because they assumed a small number of loci ($\sim 10^4$), under which only VSDMs with $s > 10^{-4}$ can make a cumulative impact. However, if individual nucleotide sites are treated as loci, their number can exceed 10^9. Crow (1972, 1993) argued that the impact of VSDMs "may be well neutralized by the extinction of small populations accumulating too many such mutants" (p. 14). However, the effective population size, N_e, of the whole species can be small enough to permit accumulation of VSDMs, possibly causing its complete extinction. Using data on DNA sequences, Tachida (1990) concluded that VSDMs impairing only one function of DNA—its interaction with nucleosomes—may lead to too high a mutation load.

0022–5193/95/160583+12 $12.00/0

© 1995 Academic Press Limited

The impact of VSDMs on survival of small sexual populations has been considered by Gabriel *et al*. (1991), Lande (1994) and Lynch *et al*. (1994). They assumed independent selection against different mutations and concluded that after a drop in N_e to a new value, the VSDMs begin to accumulate after $\sim 4N_e$ generations and then can rapidly drive the population to extinction when $N_e < 100-1000$. Mutations with $s \sim 1/N_e$ are the most dangerous, because those with higher s have no chances to be fixed, while those with lower s do less damage (Lande, 1994). Simultaneous selection against many mutations can lead to the further decline of N_e and facilitate extinction (Li, 1987; Lynch *et al*., 1994). These results are relevant to the problems of conservation of the populations that experience a recent drop in size and are thus far from equilibrium.

Here, I address the complementary problem: the impact of VSDMs for equilibrium fitness. This is justified when the N_e of a species made it vulnerable to VSDMs over a very long time. First, I present heuristic arguments, and supplement them with a simple one-locus model. Then, I interpret the results in terms of the whole genome and show, in agreement with Tachida (1990), that VSDMs can cause too high a mutation load even when $N_e \sim 10^6$–10^7. After this, the data on the relevant parameters in nature is reviewed, showing that the conditions under which the load may be paradoxically high are quite realistic. Finally, possible resolutions of this paradox are discussed.

Heuristic Arguments

Consider a locus A with two alleles A and a. The mutation rate from A to a is μ, and from a to A is v. Mating is random and the effective population size is N_e. Selection occurs in the diplophase, and the genotypes AA, Aa, and aa have fitnesses 1, 1-s, and 1-2s, respectively, where $0 < s < 0.5$ is the selection coefficient against a.

I shall treat nucleotide sites as loci and consider substitutions only. Then, A is the best nucleotide at some site, while a denotes collectively three other nucleotides (Kimura, 1983: 197; Li, 1987: 338), and $\mu \approx v$ (with all substitutions equally frequent, $\mu = 3v$). Of course, consideration of deletions and insertions would imply $\mu \gg v$.

If $s \ll \mu$, mutation is a stronger force than selection and the expected value of the frequency of a, x, is always close to the mutational equilibrium $\mu/(\mu+v) \approx 0.5$ (the variance of x declines when N_e grows). Thus, because a decreases fitness by s, the mutation load (see Crow, 1970) is $\sim s$ under any μ and N_e (Kimura *et al*., 1963).

In contrast, if $s \gg \mu$, the load depends on N_e. When N_e is large enough ($4N_e s \gg 1$), random drift is unimportant and x is always low ($\sim \mu/s$), thus deviating from the mutational equilibrium. The maintenance of this deviation requires that each generation almost all alleles that mutated from A to a are eliminated, because a and thus mutations from a to A, are rare. Thus, the load is $\sim \mu$ (see Crow, 1970; Burger & Hofbauer, 1994). However, when $4N_e s \ll 1$ the expected x grows to $\sim \mu/(\mu+v)$, and the load becomes $\sim s$, so that a decline in N_e leads to a higher load (Fig. 1). The transition between the asymptotics of large and small $4N_e s$ essentially occurs while it changes within one order of magnitude (Kimura *et al*., 1963, their figure 2).

Let us now consider the genome consisting of G nucleotides, assuming that μ and s are the same at all G sites. If the genomic deleterious mutation rate $U = G\mu$ exceeds one, the total mutation load can be too high even when drift is not important. This deterministic mutation load paradox can be resolved if selection is epistatic, which leads to a smaller load under a given mutation rate (Kimura & Maruyama, 1966; Crow, 1970; Shnol & Kondrashov, 1994).

Here, I consider the stochastic mutation load paradox, which may occur when drift increases the load (Fig. 1). If the total load is related to the sum of selection coefficients at individual sites (number of lethal equivalents, Crow, 1970), which is true if selection acts on them independently, it may be too high ($\sim Gs \gg 1$) in a small population, but not in a large one ($\sim G\mu < 1$). I call deleterious mutations

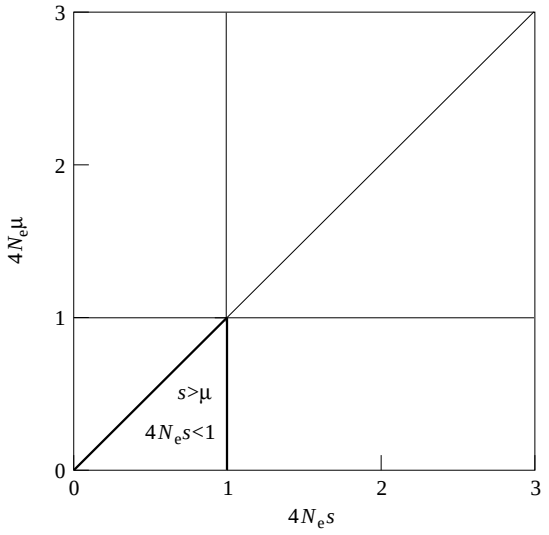

FIG. 1. The region in the space of parameters s, μ, and N_e where the stochastic mutation load paradox is possible is bounded by the bold line. High frequencies of mutant alleles require $4N_e s < 1$, while the increase of the load due to random drift requires $s > \mu$. Thus $4N_e\mu < 1$ is implied.

TABLE 1
The numerically calculated $M[\phi]$ as a function of $4N_e s$ and $4N_e\mu$ when $\mu = v$

	$4N_e s$						
$4N_e\mu$	1000	100	10	1	0.1	0.01	0.001
1000	0.382	0.488	0.499	0.500	0.500	0.500	0.500
100	0.090	0.382	0.488	0.499	0.500	0.500	0.500
10	0.008	0.091	0.387	0.488	0.499	0.500	0.500
1	0.001	0.004	0.014	0.319	0.481	0.498	0.500
0.1	0.000	0.001	0.002	0.271	0.475	0.498	0.500
0.01	0.000	0.001	0.001	0.270	0.475	0.498	0.500
Equation (3)	0.000	0.000	0.000	0.269	0.475	0.498	0.500

with $s < 1/G$ and $s < 1/4N_e$ "genomically neutral" and "populationally neutral", respectively. Even if genomically neutral nucleotides occupy all the genome, their cumulative effect is still small. Thus, because only a populationally neutral mutation can reach a high frequency, random drift causes no new problems if $1/G > 1/4N_e$. In contrast, if $G > 4N_e$, there is a dangerous range of the values of s, $1/G < s < 1/4N_e$. A mutation with such s is genomically deleterious but populationally neutral. It can reach a high frequency, and the total load caused by such mutations may be too high.

In many species (see below) $G \approx 10^9$ and $N_e \approx 10^4$. Then, if at 10% of all nucleotides the values of s are below 10^{-5}, with the average $\sim 10^{-6}$, it apparently implies ~ 100 ($10\% \times 10^9 \times 10^{-6}$) lethal equivalents per genome, which is reflected in the title of this paper.

Formal Model of a Single Locus

In the situation described above, $\phi(x)$, the probability that at a given moment of time x is between $x - \delta x/2$ and $x + \delta x/2$, is:

$$\phi(x) = C\, e^{-4N_e s x} x^{4N_e\mu - 1}(1-x)^{4N_e v - 1} \qquad (1)$$

where C can be found from the condition

$$\int_0^1 \phi(x)\, dx = 1$$

(Crow & Kimura, 1970, eqs 9.3.3 and 9.1.4). The expected frequency of x

$$M[\phi] = \int_0^1 x\phi(x)\, dx \qquad (2)$$

can be generally expressed only as the sum of an infinite series (Li, 1987). However, if μ and v are both small,

$$M[\phi] \approx \frac{\mu/v}{\mu/v + e^{4N_e s}}. \qquad (3)$$

Thus, in this low mutation rate limit $M[\phi]$ approaches zero when $s \gg 1/4N_e$, and $\mu/(\mu+v)$ when $s \ll 1/4N_e$. Numerical results (Table 1) show that $M[\phi]$ is always close to (3) when $4N_e\mu$, $4N_e v < 1$, so that for our purposes (3) is sufficient (Fig. 1). Here (3) is derived directly from (1) (see Appendix; compare with Li, 1987, eq. 8; Zeng *et al.*, 1989; Bulmer, 1991, eq. 6).

Let us now calculate $V[n]$, the expected populational variance of n, the number of deleterious alleles at locus A per diploid individual. Under a given x the variance of n is, assuming Hardy–Weinberg proportions, $2x(1-x)$. Thus

$$V[n] = 2\int_0^1 x(1-x)\phi(x)\, dx$$

$$= 2\left(M[\phi] - \int_0^1 x^2\phi(x)\, dx \right). \qquad (4)$$

Of course, $V[n]$ is different from

$$V[\phi] = \int_0^1 x^2\phi(x)\, dx - M[\phi]^2,$$

the variance of $\phi(x)$. With $4N_e\mu$, $4N_e v \ll 1$, when $\phi(x)$ is U-shaped because most of the time either allele is fixed or close to fixation, $V[\phi]$ is high while $V[n]$ is small (Crow & Kimura, 1970, eq. 9.2.7, fig. 9.3.1).

When μ and v tend to zero, $V[n]$ also approaches zero. If $r = \mu/v$ remains invariant, the rate of this approach, relative to the geometric mean of $4N_e\mu$ and $4N_e v$, is (see Appendix)

$$R = \frac{\partial V[n]}{\partial(4N_e\sqrt{\mu v})} = \frac{1}{\sqrt{r}}\frac{\partial V[n]}{\partial(4N_e v)}$$

$$= \frac{2\sqrt{\mu/v}(1 - e^{-4N_e s})}{4N_e s(1 + (\mu/v)\, e^{-4N_e s})}. \qquad (5)$$

Very Slightly Deleterious Mutations in the Whole Genome

Assume that at all G sites of the genome μ and v are the same and all the processes at them are independent. If $q(s)$ is the fraction of sites with selection coefficients against a between $s-\delta s/2$ and $s+\delta s/2$, the proportion of populationally neutral sites in the genome is

$$f=\int_0^{1/4N_e} q(s)\,\mathrm{d}s \tag{6}$$

and the average selection coefficient at such sites is

$$\bar{s}=f^{-1}\int_0^{1/4N_e} sq(s)\,\mathrm{d}s. \tag{7}$$

Let us define contamination of an individual diploid genotype D as the sum of selection coefficients against all its deleterious nucleotides. Thus, D is also a number of lethal equivalents, where a lethal equivalent is a group of mutations that, if dispersed in different individuals chosen at random from the population, would diminish their expected number of offspring by one (compare with Crow, 1970: 135). The fitness of an individual w depends on D.

Because $M[\phi]\approx 0.5$ with $\mu\approx v$ and $s<1/4N_e$ [eqn (3)], the average of the populational distribution of D, $\psi(D)$, is

$$M[\psi]=2G\int_0^{1/4N_e} M[\phi]sq(s)\,\mathrm{d}s\approx Gf\bar{s}. \tag{8}$$

and, because $R\approx 1$ [eqn (5)], the variance of $\psi(D)$, ignoring linkage disequilibrium and assuming $4N_e\mu$, $4N_e v\ll 1$, is:

$$V[\psi]=G\int_0^{1/4N_e} 4N_e\sqrt{\mu v}Rsq(s)\,\mathrm{d}s\approx 4N_e\mu Gf\bar{s}. \tag{9}$$

Because $\sqrt{V[\psi]}\ll M[\psi]$, the mean population fitness is close to $w(M[\psi])$, if $w(D)$ changes slowly. Thus, without epistasis, i.e. under exponential selection (Shnol & Kondrashov, 1993) $w(D)=\mathrm{e}^{-D}$ (we can assume $w(0)=1$ without loss of generality, while the coefficient with D must be one because of the definition of D) the load is

$$L=1-\mathrm{e}^{-M[\psi]}=1-\mathrm{e}^{-Gf\bar{s}}. \tag{10}$$

Such a load is tolerable only if

$$Gf\bar{s}\lesssim 1, \tag{11}$$

i.e. when the average contamination of individuals does not greatly exceed 1. This is always the case if $G<4N_e$, so that $\bar{s}<1/G$. In contrast, with $G>4N_e$ the load can be too close to 1.

To estimate the minimal load possible under any $w(D)$, suppose that $\psi(D)$ is Gaussian, which is reasonable because D consists of many independent (assuming linkage equilibrium) contributions. Than, the selection coefficient against a small group of individuals with the distribution of D slightly shifted by k to the right, relative to the rest of the population, is (Kimura & Crow, 1978, eq. 16a)

$$s_k=k\frac{-\Delta}{V[\psi]}, \tag{12}$$

where Δ is the selection differential of D. Such a group can consist of those carrying an independently distributed rare allele that contributes k into D. Because of the definition of D, $s_k=k$, so that $V[\psi]=-\Delta$. Even under the most efficient truncation selection, which is the extreme case of synergistic epistasis, the load is too high if $|\Delta|/\sqrt{V[\psi]}>2$ (see Shnol & Kondrashov, 1994; Kondrashov, 1994), which becomes $\sqrt{V[\psi]}>2$. Thus, according to (9) the load is tolerable when

$$N_e\mu Gf\bar{s}<1. \tag{13}$$

Therefore, the stochastic mutation load paradox appears readily under exponential selection due to violation of (11) (Tachida, 1990), but is less likely to appear under synergistic epistasis where violation of (13) is necessary, because we have assumed small $N_e\mu$.

Data on Natural Populations

We have seen that under some of G, μ, N_e, and $q(s)$ random drift can seriously aggravate the impact of VSDMs on fitness in an equilibrium population. While the stochastic mutation load paradox is hardly applicable to unicellular organisms with their huge populations and relatively small genomes, in many multicellular organisms $G\gg N_e$. Let us consider each parameter separately.

THE GENOME SIZE G

The genome sizes can be rather variable even within a genus. For each taxon we should look for a species with the smallest genome, because it probably contains the minimal fraction of genomically neutral "junk" DNA. The minimal known values of G are 70×10^6 in flowering plants (*Arabidopsis thaliana*, Meierowitz, 1987), 140×10^6 in insects (*Drosophila melanogaster*, see Lewin, 1994, Ch. 22), and 400×10^6 in fishes (*Fugu rubripes*, Brenner *et al*., 1993). Birds (1200×10^6 in

Gallus domesticus) and mammals (3300×10^6 in *Homo sapiens*) (see Lewin, 1994, Ch. 22) have rather uniform G's. Over 50% of DNA in all these species is unique.

THE PER NUCLEOTIDE MUTATION RATE μ

Various methods indicate that in mammalian nuclear DNA $\mu \approx 10^{-8}$ (see Britten, 1986; Kondrashov, 1988; Takakata, 1993; Mohrenweiser, 1994), and is probably $\sim 10^{-7}$ (Takahata, 1993) or even higher (Lundstrom *et al*., 1992) in mitochondria. Although at some nuclear sites μ can be much higher than 10^{-8} (Jeffreys *et al*., 1991; Bissler *et al*., 1994; Rousseau *et al*., 1994), the background mutation rate probably stays within the same order of magnitude along the genome.

THE EFFECTIVE POPULATION SIZE N_e

The properties of equilibrium under $\mu \approx 10^{-8}$ depend on the long-term N_e of the whole species. In a local population N_e is usually substantially lower than N_T, the total number of individuals (see Crow & Kimura, 1970; Nei, 1987; Harris & Allendorf, 1989; Allendorf *et al*., 1991). Spatial structure complicates the concept of N_e (Sugg & Chesser, 1994). Variance N_e increases with constant spatial structure (Maruyama, 1972, 1977; Nei & Takahata, 1993), while with local populations disappearing and reappearing it decreases drastically (Maruyama & Kimura, 1980; Barton, 1993; Whitlock, 1994). However, we are interested in a different N_e, which characterizes efficiency of selection. I am obliged to Nick Barton for the following explanation.

Spatial structure always reduces the efficiency of selection, as long as different local populations have different allele frequencies, because selection operates slower when an allele is closer to fixation. Even if N_T of the species is infinite, spatial subdivision can lead to fixation of the allele a when $v = 0$. In this case with $s \gg \mu$ the equilibrium frequency of a is $\mu/[s(1\text{-}F_{ST})]$, instead of μ/s without spatial structure, where Wright's fixation index F_{ST} measures genetic differentiation among local population (see Crow & Kimura, 1970). Thus, we can expect N_e (in the sense relevant here) of a species to be smaller than its N_T. Of course, N_e may exceed N_T temporarily if N_T is small because of unusual environmental conditions or anthropogenic impact, which is frequently the case for species with small N_T.

In some species with large body size, especially those occupying the upper positions in the trophic chains, even the highest N_T possible when the range and the density of a species are maximal, is low. The breeding density of the birds of prey with female body-weight of 3 kg or more is about one pair per 100 km^2 (Newton, 1979, figure 10), and the ranges of such species imply $N_T \sim 10^4$–10^5, and may be even less. Large baleen whales and the larger toothed whales had $N_T \approx 10^5$ before the impact of whaling (Ridgway & Harrison, 1985, 1989), despite their huge ranges. The total number of grizzly bears before the arrival of Europeans is estimated as 10^5 (Allendorf *et al*., 1991). In a species with a narrow range N_T can be even lower. For example, in *Lathimeria chalumnae* $N_T < 10^4$ (Thomson, 1991: 220), and apparently it has not been much higher for quite a long time. Other examples can be found in Nei & Graur (1984).

To interpret the data on N_T in terms of N_e we need to know the history of N_T (Slatkin, 1987; Slatkin & Barton, 1989; Hudson *et al*., 1992) and F_{ST}. In many species F_{ST} is low (Georgiadis *et al*., 1993; Takahata, 1993), while in other, sometimes closely related species F_{ST} is high (Cronin *et al*., 1991; Routman, 1993), suggesting $N_e < N_T$.

Alternatively, the N_e of a species (more precisely, $4N_e\mu$) can be estimated directly from the data on its neutral genetic variability, regardless of N_T. Data on protein electromorphs (see Nevo *et al*., 1984) suggest that in most vertebrates $N_e \approx 10^4$–10^5, while in invertebrates it is 10–100 times higher (Nei & Graur, 1984). These data frequently indicate that $N_e \ll N_T$. Many not very large mammals, perhaps with $N_T \geqslant 10^6$, have low variability implying $N_e \leqslant 10^4$ (Allendorf *et al*., 1979; Simonsen, 1982; Simonsen *et al*., 1982*a*, *b*).

However, it is still not clear to what extent electromorphs are neutral, and the data on the variability at the DNA level are preferable from this perspective (Karl & Avise, 1992). Several formal methods are available to infer N_e from such data (see Fu, 1994). For many species, such estimates yield $N_e \sim 10^4$, e.g. for the polar bear (Cronin *et al*., 1991), moose (Cronin, 1992), African elephant (Georgiadis *et al*., 1994), red-winged blackbird, American eel, and hardhead catfish (Avise *et al*., 1988; Table 1), and several other fish species (Gold, 1993, his Table 6), while N_T was much higher in some cases. This may be caused by fluctuations of N_T, by spatial structure, and by the impact of selection (Bulmer, 1991; Begun & Aquadro, 1992; B. Charlesworth *et al*., 1993). In *Homo sapiens*, where now F_{ST} is low, $N_e \approx 10^4$ in the last 1 Myr (Takahata, 1993).

We may conclude that both ecological and genetical evidence suggest that in many large and even not so large vertebrates, even those occupying (or have occupied before the anthropogenic impact) wide ranges, the long-term N_e of the whole species is $\sim 10^4$. Species with narrow ranges may have even smaller N_e's. Many plants also have very low N_e's due to inbreeding (Breiman *et al*., 1991; Husband & Barrett,

1992; Parra *et al*., 1993; Eguiarte *et al*., 1993), although it is unclear for how long they remain so low.

THE DISTRIBUTION OF SELECTION COEFFICIENTS $q(s)$

In nature, most deleterious mutations have only slight effects and are semi-dominant (see Crow & Simmons, 1983). Even for null-alleles of the loci studied by electrophoresis, s is only $\sim 10^{-3}$ (see Gillespie, 1991: 60; Harada *et al*., 1993). However, the data on DNA hybridization suggest that in mammals nucleotide substitutions are populationally deleterious at least at $\sim 10\%$ of the unique DNA (Britten, 1986), because $\sim 10\%$ of such DNA hybridizes between species that diverged more than μ^{-1} generations ago. If the large-scale heterogeneity in the DNA composition in mammals and birds is maintained by selection (isochores, see Bernardi, 1993), this would also suggest that $s > 1/4N_e$ for very many substitutions.

If so, what is left for sites having selection coefficients within the dangerous range? To measure s directly for such sites is difficult, because suboptimal nucleotides are frequent there. When s declines, the transition from low to high frequency of such nucleotides is rather abrupt. Thus, an intermediate frequency of apparently suboptimal nucleotides, although explainable by the values of $4N_e s$ within a narrow range (0.4–0.8, Tachida, 1990) more realistically implies heterogeneity of selection coefficients.

According to (3), f can be indirectly measured as twice the frequency of suboptimal nucleotides in the genome, which can be estimated as follows. Consider two lineages originating from a common ancestor with the N_e values N_1 and N_2 ($N_1 > N_2$). Then, ignoring beneficial mutations, we can estimate the fraction of nucleotide sites where $1/4N_1 < s < 1/4N_2$ by comparing the substitution rates in these lineages, because in the first lineage such sites are always occupied by the best nucleotides and thus do not evolve, while in the second lineage they are populationally neutral and thus evolve almost as fast as neutral ones (Ohta, 1992, 1993*a*).

This was recently done by Ohta (1993*b*), who compared the rates of synonymous and replacement nucleotide substitutions in protein-coding DNA of rodents and primates. She has found that the per generation rate of replacement substitutions was approximately two times higher in primate lineage, while synonymous substitution rates were roughly the same. She concluded that synonymous substitutions are populationally neutral in both lineages, while the fraction of populationally neutral replacement substitutions in the primate lineage is twice as large as that in the rodent lineage.

Suppose that $N_1 \approx 10^6$ (rodents) and $N_2 \approx 10^4$ (primates) (Nei & Graur, 1984). Their common ancestor probably had $N_e \approx N_1$, so that the difference between the lineages is caused by accumulation of VSDMs in primates. Then, because in the primate lineage in ~ 11000 nucleotides there were ~ 650 replacement substitutions, roughly half of them "excessive", Ohta's conclusion implies that in at least in $\sim 3\%$ of nucleotides in the protein-coding regions s is within 10^{-5}–10^{-7}. It is unclear how many nucleotides have $s < 10^{-7}$, but it is probably a lot, because in rodents there also were 1000 replacement substitutions, although some of them may be beneficial.

To interpret these data in terms of the whole genome, note that in mammals $\sim 600 \times 10^6$ nucleotides are transcribed in the brain (Takahashi, 1992), and $\sim 1000 \times 10^6$ nucleotides are transcribed in at least one tissue (Evtushenko *et al*., 1989). Assuming that only 10% of length of the nuclear RNA is later translated, this implies that the total length of protein-coding regions P is $\sim 100 \times 10^6$. This is consistent with the data on yeast, where $G = 14 \times 10^6$ and $> 50\%$ of DNA in chromosome III is coding, implying $P \sim 10 \times 10^6$ (Oliver *et al*., 1992) and *C. elegans*, where $G \approx 10^8$ and $\sim 30\%$ of a long region of chromosome III is coding, implying $P \sim 30 \times 10^6$ (Wilson *et al*., 1994).

Thus, Ohta's data imply that the genome of a primate carries at least 3×10^6 suboptimal nucleotides only in the protein-coding regions (and perhaps much more: see Eastreal & Collet, 1994). Of course, sites with dangerous values of s also occur in non-translated and even in non-transcribed DNA, because the pattern of selection in the adjacent coding and non-coding regions may be similar (Li & Salter, 1991). A similar analysis would be interesting for non-transcribed DNA, where pseudogenes can serve as the neutral markers, while changes in 5′-upstream regulatory regions can be analogous to replacement substitutions. It may also be worth comparing the patterns of substitutions in the pairs of lineages that underwent profound phenotypic changes, and those that remained essentially unchanged for a long time.

Discussion

The data strongly suggest that, at least in many vertebrates, $G \gg N_e$. Thus, the dangerous range of selection coefficients $1/G < s < 1/4N_e$ can be wide. The values of μ are low enough to make the condition $s \gg \mu$ true for most of this range, so that random drift can, indeed, substantially increase the mutation load. In contrast to the deterministic case ($s > 1/4N_e$, see Crow, 1970; Kondrashov, 1988, 1993), the load can become excessive even when $U < 1$.

Traditionally (see Ohta, 1992) mutations with $s < 1/4N_e$ (sometimes other similar thresholds are used) are called very slightly deleterious. However, in an equilibrium population all four nucleotides have similar expected frequencies at a site with such s, so that a mutation has similar chances of being either very slightly deleterious or very slightly beneficial. Substitutions (fixations or quasifixations of the best or a suboptimal nucleotide) at such a site occur almost at the rate characteristic of neutral sites, on average once in μ generations (Tachida, 1990, figure 5). Of course, deleterious mutations can outnumber beneficial temporarily, after a drop of N_e (Lande, 1994; Lynch *et al*., 1994).

Contamination by VSDMs is related to Muller's ratchet (see Lynch *et al*., 1993) because both are stochastic phenomena. However, the ratchet operates when individual alleles are still rare, and thus requires obligate or almost obligate asexuality or selfing (see D. Charlesworth *et al*., 1993; Lynch *et al*., 1993). In contrast, the accumulation of VSDMs is possible under sex and outcrossing, although asexuality, which diminishes N_e due to more intensive hitch-hiking, can facilitate it (Charlesworth, 1991; B. Charlesworth *et al*., 1993; Rice, 1994; Lynch *et al*., 1994).

The principle question is whether—as my analysis apparently suggests—contamination by VSDMs implies an excessive load, leading to the stochastic mutation load paradox. I shall first review the simplifications used in my analysis. Then the reality of this paradox will be assessed. Finally, possible resolutions will be discussed.

SIMPLIFYING ASSUMPTIONS

I have assumed $\mu \approx v$. In fact, $\mu \geqslant v$. Even if only one suboptimal nucleotide is populationally neutral, while the other two have $s > 1/4N_e$ and have to be ignored, $\mu = v$ if mutation rates are symmetric. Thus eqn (8) underestimates the expected contamination.

I have ignored insertions and deletions, which is justified for protein-coding DNA, where they are usually significantly deleterious. However, in other regions (see Gillespie, 1991: 78) they may be common and populationally neutral. First, with insertions and deletions the rate of forward mutation greatly exceeds the backward rate, leading to a higher equilibrium contamination. Second, they make it impossible to consider different sites separately, increasing the number of trajectories in the space of sequences (Gillespie, 1984).

Let us arrange all the sequences linearly from zero to infinity, and assume that only mutations between the adjacent states have non-negligible frequencies (stepwise mutation model, Kimura, 1983: 229). The number of a state indicates the minimal number of mutations that separates it from the best one, and the fitness decreases with this number. The frequency distribution of the state number can reach an equilibrium. However, if mutations that increase the number of a state are more frequent than backward mutations, weak selection in a finite population may be unable to prevent the unlimited increase of the average state number and, thus, the unlimited decline of fitness. A similar phenomenon is possible in infinite populations under inefficient selection with diminishing returns epistasis (Kimura & Maruyama, 1966).

It is unclear how natural populations can avoid this. One possibility is that after some deviation from the optimum each next mutation becomes very deleterious (synergistic epistasis, see below). Anyway, the assumption that there is an equilibrium contamination can only underestimate the impact of VSDMs.

Finally, I have extrapolated the results from an individual site to the whole genome. There is a consensus that this leads to underestimation of the genome contamination (Li, 1987; Bulmer, 1991; B. Charlesworth *et al*., 1993; Lynch *et al*., 1994), because selection processes at different sites interfere with each other. However, this can be probably taken care of by assuming a lower N_e (Bulmer, 1991). Note, that the values of N_e measured from the data on molecular variability take the results of selection into account.

IS THE STOCHASTIC MUTATION LOAD PARADOX REAL?

According to Kimura (1983: 248) VSDMs do not cause any problem because (i) they accumulate very slowly and (ii) their impact can be easily counterbalanced by rare fixations of beneficial alleles. I do not think that this is correct.

High contamination by VSDMs is reached after N_e of a lineage remains much smaller than G during $\sim \mu^{-1} \approx 10^8$ generations. This may be the case in some vertebrates. In addition, if after a drop of N_e the expected equilibrium contamination is, say, 100, VSDMs may become important much sooner, $\sim 10^6$ generations after the drop. The total mutation rate in mammals is ~ 100 events per genome. If 10% of them are VSDMs with the average selection coefficient 10^{-6} (implying $N_e \approx 10^5$), they cause the decline of fitness by $\sim 10^{-5}$ per generation (if initially all nucleotides were best). This decline will become important $\sim 10^5$ generations after the drop of N_e.

A beneficial mutation is possible only in a site occupied by a suboptimal nucleotide. This implies that either a VSDM was previously fixed at this site (in this case the beneficial mutation can have only a very small advantage) or that the environment has changed (then

the mutation can be significantly beneficial). Very slightly beneficial mutations are part of our analysis: they appear when VSDMs are frequent and prevent them from fixation, while changes of the environment, either in time or in space, can only add a component to the genetic load (lag-load or migration load, respectively, see Crow, 1970; Stenseth & Maynard Smith, 1982). Thus, the load is minimal when, as it was assumed here, the selection remains invariant in time and space for a long time, so that all significantly beneficial mutation are already fixed.

POSSIBLE RESOLUTIONS AND IMPLICATIONS OF THE PARADOX

Because the stochastic mutation load paradox appears real, it requires a resolution. Five options must be mentioned.

(i) First, the properties of equilibrium may be irrelevant, because it is never reached (Crow, 1972). Chetverikov (1926, Result 7) assumed that the mutational contamination of a species increases with time, leading, perhaps, to its eventual extinction. In reality, of course, there is no reason why a new species should have a fresh start. However, the suggestion that accumulation of VSDMs in a lineage with $N_e \ll G$ acts as a time-bomb, finally destroying it, is a modification of this idea. If so, the existence of vertebrate lineages with $N_e \leqslant 10^4$–10^5 should be limited to $\sim 10^6$–10^7 generations. Perhaps, this does not contradict the existing data.

Otherwise, the parameters of natural populations may not lead to too high contaminations even if the equilibrium is reached. Because G and μ are known reasonably well, I shall consider only N_e and $q(s)$.

(ii) N_e could be higher than I assumed if the actual pools of the available genes are larger than what is usually recognized as "species". However, although interspecific hybridization (e.g. Spilliaert *et al*., 1991) and horizontal gene transfer (Clark *et al*., 1994) may occur, it is highly unlikely that they are so important.

(iii) The nucleotide sites with dangerous values of s may be rare. This implies that $q(s)$ is bimodal, with the dip in the range of $10^{-9} < s < 10^{-5}$. It is unclear why this should be the case. Perhaps, the evolution of molecular interactions could make them insensitive to VSDMs, by limiting the involved nucleotides or amino acids to those under substantial selection.

Finally, an average genome may carry many lethal equivalents of VSDMs, but the load may still be tolerable. In fact, the principal problem [see eqns (8) and (10)] is that under exponential $w(D)$ the mean fitness of a population with $M[\psi] \gg 1$ is too low, compared to $w(0)$. Two properties of selection can change this conclusion.

(iv) $w(D)$ may involve synergistic epistasis (Tachida, 1990). The condition $-\Delta = V[\psi]$ [eqn (12)] carries information only about the rate of decline of $w(D)$ within the narrow range of contaminations actually present in the population, say $M[\psi] \pm 3\sqrt{V[\psi]}$. However, in general we cannot extrapolate $w(D)$ outside this range and, if $w(D)$ changes slowly where D is small, $w(0)$ may be close to $w(M[\psi] - 3\sqrt{V[\psi]})$.

(v) Selection against VSDMs may be soft (see Kondrashov, 1995). Then in the population with a lower $M[\psi]$, where a predicted high $w(0)$ becomes relevant, fitnesses of all genotypes decline.

Of course, selection can be simultaneously soft and epistatic. These possibilities are illustrated in Fig. 2. Because under both epistatic and soft selection the scaling on D axis generally changes with ψ, I use the number of deleterious alleles as an independent variable, assuming that selection coefficients are the same at all sites. Generally, soft and/or synergistic selection make it possible to avoid too high a load only if it is implied, as above, by extrapolating the fitness function to the best possible genotype (see Lewontin, 1974, figure 17 for a similar analysis of multiple loci with overdominance).

In contrast, if too high a load is implied by a large ratio of selection differential to the standard deviation of the trait, soft selection is irrelevant, while synergistic epistasis can make the load acceptable only if this ratio is below ~ 2 [eqn (13)]. Because $\mu G f = U$ and $N_e \bar{s} \ll 1$ (only populationally neutral mutations are considered), this requires $U \gg 1$, when the deterministic mutation load paradox (see Kondrashov, 1995) is probably more important. However, the analogous problem may be relevant for overdominance.

If all individuals indeed carry many lethal equivalents of VSDMs, the perfect mutation-free genotype is rather different from the best available one, even if this does not directly translate into high mutation load. One possible consequence of this is an advantage to sex, if transition to asexuality significantly diminishes the N_e (this idea was independently proposed and is now developed by Joel Peck). However, this advantage can hardly provide a short-term protection to sex (see Kondrashov, 1993): after a drop in N_e a mutation that became populationally neutral will require $\sim 4N_e$ generation to be fixed, and the new equilibrium will be reached only after $\sim 1/\mu$ generations (Lande, 1994; Lynch *et al*., 1994). In contrast, this effect may, together with the Muller's ratchet (D. Charlesworth *et al*., 1993; Lynch *et al*., 1993), limit the lifespan of the asexual forms.

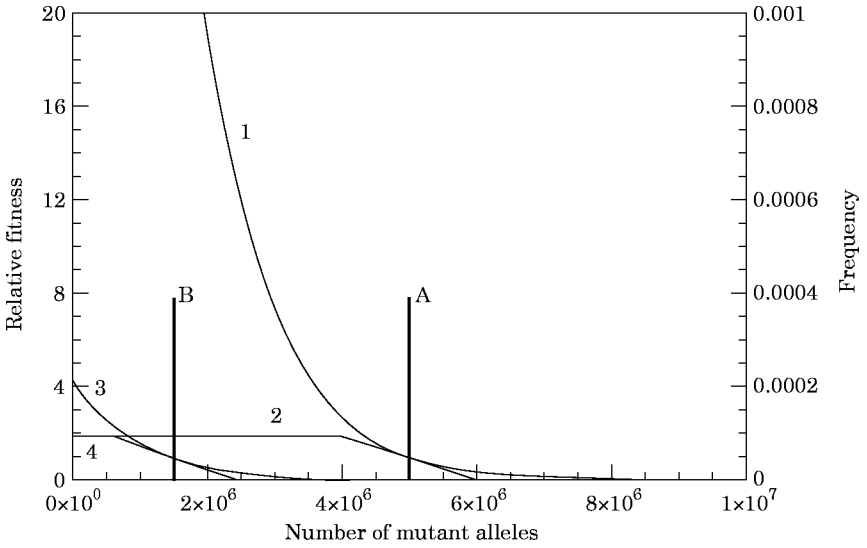

FIG. 2. Possible resolutions of the stochastic mutation load paradox by postulating epistatic and/or soft selection. I assume 10^7 sites with identical parameters, such that, according to eqn (3), the average of the genomic number of mutant alleles, x, is either 5×10^6 (A, a "natural" population) or 1.5×10^6 (B, a relatively mutation-free population). Under all realistic parameters the variance of x is too small to be visible on the figure. Four fitness functions $w(x)$ are considered, and the fitness of a genotype relative to the population mean is plotted. When population A is under exponential selection $w(x) = e^{sx}$, with selection coefficient at a site $s = 10^{-6}$ (plot 1), the ratio of the mean population fitness over $w(0)$, W, is <0.01. However, under epistatic $w(x)$ which causes practically the same s (plot 2), W in A is 0.5. Under exponential soft selection (plot 3) in population B, W is ~ 0.2, although s remains the same. If selection is both epistatic and soft, plot 4 is the fitness function in B and $W = 0.5$.

I thank Nick Barton, Ben Normack, Sally Otto, Joel Peck, and the whole Chip Aquadro team for many helpful comments.

REFERENCES

ALLENDORF, F. W., CHRISTIANSEN, F. B., DOBSON, T., EANES, W. F. & FRYDENBERG, O. (1979). Electrophoretic variation in large mammals I. The polar bear, *Thalarctos maritimus*. *Hereditas* **91,** 19–22.

ALLENDORF, F. W., HARRIS, R. B. & METZGAR, L. H. (1991). Estimation of effective population size of grizzly bears by computer simulation. In: *The Unity of Evolutionary Biology* Vol. II (Dudley, E. C., ed.) pp. 650–654. Portland, OR: Dioscorides Press.

AVISE, J. C., BALL, R. M. & ARNOLD, J. (1988). Current versus historical population sizes in vertebrate species with high gene flow: a comparison based on mitochondrial DNA lineages and inbreeding theory for neutral mutations. *Molec. Biol. Evol.* **5,** 331–344.

BARTON, N. H. (1993). The probability of fixation of a favoured allele in a subdivided population. *Genet. Res.* **62,** 149–157.

BEGUN, D. J. & AQUADRO, C. F. (1992). Levels of naturally occuring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature* **356,** 519–520.

BERNARDI, G. (1993). The vertebrate genome: isochores and evolution. *Molec. Biol. Evol.* **10,** 186–204.

BISSLER, J. J., CICARDI, M., DONALDSON, V. H., GATENBY, P. A., ROSEN, F. S., SHEFFER, A. L. & DAVIS, A. E. (1994). A cluster of mutations within a short triplet repeat in the C1 inhibitor gene. *Proc. natn. Acad. Sci. U.S.A.* **91,** 9622–9625.

BREIMAN, A., BOGHER, M., STERNBERG, H. & GRAUR, D. (1991). Variability and uniformity of mitochondrial DNA in populations of putative diploid ancestors of common wheat. *Theor. appl. Genet.* **82,** 201–208.

BRENNER, S., ELGAR, G., SANFORD, R., MACRAE, A., VENKATESH, B. & APARICIO, S. (1993). Characterization of the pufferfish (*Fugu*) genome as a compact model vertebrate genome. *Nature* **366,** 265–268.

BRITTEN, R. J. (1986). Rates of DNA sequence evolution differ between taxonomic groups. *Science* **231,** 1393–1398.

BULMER, M. (1991). The selection-mutation-drift theory of synonymous codon usage. *Genetics* **129,** 897–907.

BURGER, R. & HOFBAUER, J. (1994). Mutation load and mutation-selection balance in quantitative genetic traits. *J. math. Biol.* **32,** 193–218.

CHARLESWORTH, B. (1991). The evolution of sex chromosomes. *Science* **251,** 1030–1033.

CHARLESWORTH, B., MORGAN, M. T. & CHARLESWORTH, D. (1993). The effect of deleterious mutations on neutral molecular variation. *Genetics* **134,** 1289–1303.

CHARLESWORTH, D., MORGAN, M. T. & CHARLESWORTH B. (1993). Mutation accumulation in finite outbreeding and inbreeding populations. *Genet. Res.* **61,** 39–56.

CHETVERIKOV, S. S. (1926). On certain aspects of the evolutionary process from the standpoint of modern genetics. *J. expl Biol.* **A2,** 3–54 (Russian). Translation: [1961]. *Proc. Am. Phil. Soc.* **105,** 167–195.

CLARK, J. B.. MADDISON, W. P. & KIDWELL, M. G. (1994). Phylogenetic analysis supports horizontal transfer of P transposable elements. *Molec Biol. Evol.* **11,** 40–50.

CRONIN, M. A. (1992). Intraspecific variation in mitochondrial DNA of North American cervids. *J. Mammal.* **73,** 70–82.

CRONIN, M. A., AMSTRUP, S. C., GARNER, G. W. & VYSE, E. R. (1991). Interspecific and intraspecific mitochondrial DNA variation in North American bears (*Ursus*). *Can. J. Zool.* **69,** 2985–2992.

CROW, J. F. (1970). Genetic loads and the cost of natural selection. In: *Mathematical Topics in Population Genetics* (Kojima, K.-I., ed.) pp. 128–177. New York: Springer.

CROW, J. F. (1972). Darwinian and non-Darwinian evolution. *Proceedings of the sixth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 5, pp. 1–22.

CROW, J. F. (1993). Mutation, fitness, and genetic load. *Oxf. Surv. Evol. Biol.* **9,** 3–42.

CROW, J. F. & KIMURA, M. (1970). *An Introduction to Population Genetics Theory*. New York: Harper & Row.

Crow, J. F. & Simmons, M. J. (1983). The mutation load in *Drosophila*, In: *The Genetics and Biology of Drosophila*, Vol. 3c (Ashburner, M., Carson, H. L. & Thompson, J. N., Jr, eds) pp. 1–35. London: Academic Press.

Eastreal, S. & Collet, C. (1994). Consistent variation in amino-acid substitution rate, despite uniformity of mutation rate: protein evolution in mammals is not neutral. *Molec. Biol. Evol.* **11,** 643–647.

Eguiarte, L. E., Burquez, A., Rodrigguez, J., Martinez-Ramos, M., Sarukhan, J. & Pinero, D. (1993). Direct and indirect estimates of neighborhood and effective population size in a tropical plant, *Astrocaryum mexicanum*. *Evolution* **47,** 75–87.

Evtushenko, V. I., Hanson, K. P., Barabitskaya, O. V., Emelyanov, A. V., Reshetnikov, V. L. & Kozlov, A. P. (1989). An attempt to determine the maximal expression of the rat genome. *Molekul. Biol.* **23,** 663–675 (Russian).

Fu, Y.-X. (1994). A phylogenetic estimator of effective population size or mutation rate. *Genetics* **136,** 685–692.

Gabriel, W., Burger, R. & Lynch, M. (1991). Population extinction by mutation load and demographic stochasticity. pp. 49–59 In A. Seitz & V. Loeschke (eds). *Species Conservation: A Population-biological Approach*, pp. 49–59. Basel: Birkhäuser.

Georgiadis, N., Bischof, L., Templeton, A., Patton, J., Karesh, W. & Western, D. (1994). Structure and history of African elephant populations: I. Eastern and Southern Africa. *J. Hered.* **85,** 100–104.

Gillespie, J. H. (1984). Molecular evolution over the mutational landscape. *Evolution* **38,** 1116–1129.

Gillespie, J. H. (1991). *The Causes of Molecular Evolution*. New York: Oxford University Press.

Gold, J. R., Richardson, L. R., Furman, C., King, T. L. (1993). Mitochondrial DNA differentiation and population structure in red drum (*Sciaenops ocellatus*) from the Gulf of Mexico and Atlantic Ocean. *Mar. Biol.* **116,** 175–185.

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

Harada, K., Kusakabe, S.-I., Yamazaki, T. & Mukai, T. (1993). Spontaneous mutation rates in null and band-morph mutations of enzyme loci in *Drosophila melanogaster*. *Jap. J. Genet.* **68,** 605–616.

Harris, R. B. & Allendorf, F. W. (1989). Genetically effective population size of large mammals: an assessment of estimators. *Conserv. Biol.* **3,** 181–191.

Hudson, R. R., Slatkin, M. & Maddison, W. P. (1992). Estimation of levels of gene flow from DNA sequence data. *Genetics* **132,** 583–589.

Husband, B. C. & Barrett, S. C. H. (1992). Effective population size and genetic drift in tristylous *Eichornia paniculata* (*Pontederiaceae*). *Evolution* **46,** 1875–1890.

Jeffreys, A. J., MacLeod, A., Tamaki, K., Neill, D. L. & Monckton, D. G. (1991). Minisatellite repeat coding as a digital approach to DNA typing. *Nature* **354,** 204–209.

Karl, S. A. & Avise, J. C. (1992). Balancing selection at allozyme loci in oysters: implications from nuclear RFLPs. *Science* **256,** 100–102.

Kimura, M. (1983). The neutral theory of molecular evolution. Cambridge: Cambridge University Press.

Kimura, M., & Crow, J. F. (1978). Effect of overall phenotypic selection on genetic change at individual loci. *Proc. natn. Acad. Sci. U.S.A.* **75,** 6168–6171.

Kimura, M., Maruyama, T. & Crow, J. F. (1963). The mutation load in small populations. *Genetics* **48,** 1303–1312.

Kimura, M. & Maruyama, T. (1966). The mutational load with epistatic gene interactions in fitness. *Genetics* **54,** 1337–1351.

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

Kondrashov, A. S. (1993). Classification of hypotheses on the advantage of amphimixis. *J. Hered.* **84,** 372–387.

Kondrashov, A. S. (1995). Dynamics of unconditionally deleterious mutations: Gaussian approximation and soft selection. *Genet. Res.* **65,** 113–121.

Lande, R. (1994). Risk of population extinction from fixation of new deleterious mutations. *Evolution* **48,** 1460–1469.

Lewin, B. (1994). *Genes V*. Oxford: Oxford University Press.

Lewontin, R. C. (1974). *The Genetic Basis of Evolutionary Change*. New York: Columbia University Press.

Li, W.-H. (1987). Models of nearly neutral mutations with particular implications for non-random usage of synonymous codons. *J. molec. Evol.* **24,** 337–345.

Li, W.-H. & Salter, L. A. (1991). Low nucleotide diversity in man. *Genetics* **129,** 513–523.

Lundstrom, R., Tavare, S. & Ward, R. H. (1992). Estimating substitution rates from molecular data using the coalescent. *Proc. natn. Acad. Sci. U.S.A.* **89,** 5961–5965.

Lynch, M., Burger, R., Butcher, D. & Gabriel, W. (1993). The mutational meltdown in asexual populations. *J. Hered.* **84,** 339–344.

Lynch, M., Conery, J. & Burger, R. (1994). Mutational meltdowns in sexual populations. *Evolution*, in press.

Maruyama, T. (1972). Rate of decrease of genetic variability in a two-dimensional continous population of finite size. *Genetics* **70,** 639–651.

Maruyama, T. (1977). *Stochastic Problems in Population Genetics*. Berlin: Springer.

Maruyama, T. & Kimura, M. (1980). Genetic variability and effective population size when local extinction and recolonization of subpopulations are frequent. *Proc. natn. Acad. Sci. U.S.A.* **77,** 6710–6714.

Meyerowitz, E. M. (1987). *Arabidopsis thaliana*. *A. Rev. Genet.* **21,** 93–111.

Mohrenweiser, H. (1994). Impact of the molecular spectrum of mutational lesions on estimates of germinal gene-mutation rates. *Mutat. Res.* **304,** 119–137.

Nei, M. (1987). *Molecular Evolutionary Genetics*. New York: Columbia University Press.

Nei, M. & Graur, D. (1984). Extent of protein polymorphism and the neutral mutation theory. *Evol. Biol.* **17,** 73–118.

Nei, M. & Takahata, N. (1993). Effective population size, genetic diversity, and coalescence time in subdivided populations. *J. Molec. Evol.* **37,** 240–244.

Nevo, E., Beiles, A. & Ben-Shlomo, R. (1984). The evolutionary significance of genetic diversity: ecological, demographic and life history correlates. *Lecture Notes in Biomathematics* **53,** 13–213.

Newton, I. (1979). *Population Ecology of Raptors* Berkhamsted: T & A D Poyser.

Ohta, T. (1992). The nearly neutral theory of molecular evolution. *A. Rev. Ecol. Syst.* **23,** 263–286.

Ohta, T. (1993*a*). Amino acid substitution at the Adh locus of *Drosphila* is facilitated by small population size. *Proc. natn. Acad. Sci. U.S.A.* **90,** 4548–4551.

Ohta, T. (1993*b*). An examination of the generation-time effect on molecular evolution. *Proc. natn. Acad. Sci. U.S.A.* **90,** 10676–10680.

Oliver, S. G. et al. (1992). The complete DNA sequence of yeast chromosome III. *Nature* **357,** 38–46.

Parra, V., Vargas, C. F. & Eguiarte, L. E. (1993). Reproductive biology, pollen and seed dispersal, and neighborhood size in the hummingbird-pollinated *Echeveria gibbiflora* (*Crassulaceae*). *Am. J. Bot.* **80,** 153–159.

Rice, W. R. (1994). Degeneration of a nonrecombining chromosome. *Science* **263,** 230–232.

Ridgway, S. H. & Harrison, R. (eds) (1985). *Handbook of Marine Mammals*, Vol. 3. *The Sirenians and Baleen Whales*. London: Academic Press.

Ridgway, S. H. & Harrison, R. (eds) (1989). *Handbook of Marine Mammals*, Vol. 4. *River Dolphins and the Larger Toothed Whales*. London: Academic Press.

Rousseau, F., Bonaventure, J., Legeai-Mallet, L., Pelet, A., Rozet, J.-M., Maroteaux, P., Le Merrer, M. & Munnich, A. (1994). Mutations in the gene encoding fibroblast growth factor receptor-3 in achondroplasia. *Nature* **371,** 252–254.

ROUTMAN, E. (1993). Population structure and genetic diversity of metamorphic and paedomorphic populations of the tiger salamander, *Ambystoma tigrinum*. *J. Evol. Biol.* **6,** 329–357.
SHNOL, E. E. & KONDRASHOV, A. S. (1993). Effect of selection on the phenotypic variance. *Genetics* **134,** 995–996.
SHNOL, E. E. & KONDRASHOV, A. S. (1994). Some relations between different characteristics of selection. *J. math. Biol.* **32,** 835–840.
SIMONSEN, V. (1982). Electrophoretic variation in large mammals II. The red fox, *Vulpes vulpes*, the stoat *Mustela erminea*, the weasel, *Mustela nivalis*, the pole cat, *Mustela putorius*, the pine marten, *Martes martes*, the beech marten, *Martes foina*, and the badger, *Meles meles*. *Hereditas* **96,** 299–305.
SIMONSEN, V., ALLENDORF, F. W., EANES, W. F. & KAPEL, F. O. (1982). Electrophoretic variation in large mammals III. The ringed seal, *Pusa hispida*, the harp seal, *Pagophilus groenlandicas*, and the hooded seal, *Cystophora cristata*. *Hereditas* **97,** 87–90.
SIMONSEN, V., BORU, E. W. & KRISTENSEN, T. (1982). Electrophoretic variation in large mammals IV. The atlantic walrus, *Odobenus rosmarus* (L.). *Hereditas* **97,** 91–94.
SLATKIN, M. (1987). Gene flow and the geographic structure of natural populations. *Science* **236,** 787–792.
SLATKIN, M. & BARTON, N. H. (1989). A comparison of three indirect methods for estimating average levels of gene flow. *Evolution* **43,** 1349–1368.
SPILLIAERT, R., VIKINGSSON, G., ARNASON, U., PULSDOTTIR, A., SIGURJONSSON, J. & ARNASON, A. (1991). Species hybridization between a female of blue whale *Balaenoptera musculus* and a male fin whale *Balaenoptera physalus*: molecular and morphological documentation. *J. Hered.* **82,** 269–274.
STENSETH, N. C. & MAYNARD SMITH, J. (1982). Coevolution in ecosystems: Red Queen evolution or stasis? *Evolution* **38,** 870–880.
SUGG, D. W. & CHESSER, R. K. (1994). Effective population size with multiple paternity. *Genetics* **137,** 1147–1155.
TACHIDA, H. (1990). A population genetic model of selection that maintains specific trinucleotides at a specific location. *J. molec. Evol.* **31,** 10–17.
TAKAHASHI, Y. (1992). Gene expression in cells of the central nervous system. *Prog. Neurobiol.* **38,** 523–569.
TAKAHATA, N. (1993). Allelic genealogy and human evolution. *Molec. Biol. Evol.* **10,** 2–22.
THOMSON, K. S. (1991). *Living Fossil*: *the Story of the Coelacanth*. New York: W. W. Norton & Co.
WHITLOCK, M. C. (1994). Fission and the genetic variance among populations: the changing demography of forked fungus beetle populations. *Am. Nat.* **143,** 820–829.
WILSON, R. *et al*. (1994). 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. *Nature* **368,** 32–38.
WRIGHT, S. (1929). Fisher's theory of dominance. *Am. Nat.* **63,** 274–279.
ZENG, Z.-B., TACHIDA, H. & COCKERHAM, C. C. (1989). Effects of mutation on selection limits in finite populations with multiple alleles. *Genetics* **122,** 977–984.

APPENDIX

LEMMA. Suppose that the partial derivative $f_x(x, \beta)$ of a function $f(x, \beta)$ is uniformly bounded on $[0, a] \times [0, \beta_0]$. Then,

$$\lim_{\beta \to 0+} \left[\beta \int_0^a x^{\beta-1} f(x, \beta)\, dx \right] = f(0, 0).$$

Proof. By the Mean Value Theorem, $f(x, \beta) = f(0, \beta) + f_x(\xi, \beta)x$, where $0 \leqslant \xi \leqslant x \leqslant a$. Substitution of this into the integral yields

$$A(\beta) = \beta \int_0^a x^{\beta-1} f(x, \beta)\, dx$$

$$= \beta f(0, \beta) \int_0^a x^{\beta-1}\, dx + \beta \int_0^a x^{\beta} f_x(\xi(x), \beta)\, dx$$

$$= f(0, \beta) x^{\beta}|_0^a + \beta \int_0^a x^{\beta} f_x(\xi(x), \beta)\, dx$$

$$= f(0, \beta) a^{\beta} + 0(\beta).$$

Thus, $\lim_{\beta \to 0+} A(\beta) = f(0, 0)a^0 = f(0, 0)$. QED.

This allows us to evaluate the following limits:

$$I_0 = \lim_{\gamma \to 0+} \left[\gamma \int_0^1 x^{r\gamma-1}(1-x)^{\gamma-1} e^{\alpha x}\, dx \right]$$

$$= \lim_{\gamma \to 0+} \left\{ \gamma \left[\int_0^{1/2} x^{r\gamma-1}(1-x)^{\gamma-1} e^{\alpha x}\, dx + \int_{1/2}^1 x^{r\gamma-1}(1-x)^{\gamma-1} e^{\alpha x}\, dx \right] \right\}$$

$$= \lim_{\gamma \to 0+} \left\{ \gamma \left[\int_0^{1/2} x^{r\gamma-1}(1-x)^{\gamma-1} e^{\alpha x}\, dx + \int_0^{1/2} y^{\gamma-1}(1-y)^{r\gamma-1} e^{\alpha(1-y)}\, dy \right] \right\}$$

$$= \frac{(1-0)^{\gamma-1} e^{\alpha 0}}{r} + 1^{r\gamma-1} e^{a1} = 1/r + e^{\alpha}$$

and

$$I_1 = \lim_{\gamma \to 0+} \left[\gamma \int_0^1 x^{r\gamma}(1-x)^{\gamma-1} e^{\alpha x}\, dx \right]$$

$$= 1^{r\gamma-1} e^{\alpha 1} = e^{\alpha}.$$

Where both μ and v tends to zero while $r = \mu/v = \text{const}$, $M[\phi] = I_1/I_0$, which implies eqn (3) in the main text if $\alpha = -4N_e s$.

To derive eqn (5), it is sufficient to observe that with $\gamma = 4N_e v$

$$\lim_{\gamma \to 0+} \frac{V[n]}{\gamma} = \lim_{\gamma \to 0+} \frac{2\int_0^1 x^{r\gamma}(1-x)^{\gamma} e^{\alpha x}\, dx}{\gamma \int_0^1 x^{r\gamma-1}(1-x)^{\gamma-1} e^{\alpha x}\, dx}$$

$$= \frac{2 \lim_{\gamma \to 0+} \int_0^1 x^{r\gamma}(1-x)^{\gamma} e^{\alpha x}\, dx}{I_0}$$

$$= \frac{(2/\alpha)(e^{\alpha}-1)}{1/r + e^{\alpha}}.$$

(Appendix by Alexander I. Khibnik)