

Deleterious mutation accumulation in organelle genomes

Michael Lynch & Jeffrey L. Blanchard

Department of Biology, University of Oregon, Eugene, OR 97403, USA

Key words: chloroplast genome, deleterious mutation, mitochondrial genome, Muller's ratchet, ribosomal RNA, transfer RNA

Abstract

It is well established on theoretical grounds that the accumulation of mildly deleterious mutations in nonrecombining genomes is a major extinction risk in obligately asexual populations. Sexual populations can also incur mutational deterioration in genomic regions that experience little or no recombination, i.e., autosomal regions near centromeres, Y chromosomes, and organelle genomes. Our results suggest, for a wide array of genes (transfer RNAs, ribosomal RNAs, and proteins) in a diverse collection of species (animals, plants, and fungi), an almost universal increase in the fixation probabilities of mildly deleterious mutations arising in mitochondrial and chloroplast genomes relative to those arising in the recombining nuclear genome. This enhanced width of the selective sieve in organelle genomes does not appear to be a consequence of relaxed selection, but can be explained by the decline in the efficiency of selection that results from the reduction of effective population size induced by uniparental inheritance. Because of the very low mutation rates of organelle genomes (on the order of 10^{-4} per genome per year), the reduction in fitness resulting from mutation accumulation in such genomes is a very long-term process, not likely to imperil many species on time scales of less than a million years, but perhaps playing some role in phylogenetic lineage sorting on time scales of 10 to 100 million years.

A substantial body of theory supports the hypothesis that small to moderately large populations of asexual organisms are subject to long-term deleterious mutation accumulation (Muller, 1964; Felsenstein, 1974; Haigh, 1978; Pamilo, Nei & Li., 1987; Charlesworth, Morgan & Charlesworth, 1993; Stephan, Chao & Smale, 1993; Higgs, 1994; Gessler, 1996). At least three mechanisms can promote such deterioration, and they are not necessarily independent. First, as a consequence of the recurrent introduction of new deleterious mutations each generation, classes of individuals with different fitnesses exist in any population. If the population is asexual and if, by chance, the best class either leaves no progeny in some generation or all the progeny produced by this class acquire at least one new deleterious mutation, the best class will be lost. The previously second-best class will then be advanced in status, but it will eventually suffer the same fate, and so on. In the absence of recombination, this progressive loss of fitness (Muller's ratchet) can only be avoided if back

or compensatory mutations arise at a sufficiently high rate to balance or offset the cumulative effects of deleterious mutations. Second, the build-up of repulsion linkage disequilibrium for fitness-related mutations is inevitable in a nonrecombining genome. By partially concealing genetic variation for fitness at individual loci, such disequilibrium enhances the rate of fixation of deleterious mutations and reduces the rate of fixation of beneficial mutations (Hill & Robertson, 1966; Birky & Walsh, 1988). Third, if a mutation with transient beneficial effects is pulled to fixation by positive selection, then every deleterious mutation within the genome in which the beneficial mutation first arose will be swept to fixation with it. Because the probability that the beneficial mutation will arise within a member of the current best-class is quite small, such an event will almost always reduce the mean fitness associated with background loci; so if the beneficial effects of the mutation are subsequently lost (due, for example, to a change in ecology), a substantial decline

in the fitness may be experienced. This third potential mechanism has received little attention from theoreticians.

The ultimate consequence of long-term mutational degradation of a population is extinction (Lynch & Gabriel, 1990; Lynch et al., 1993; Lynch, Conery & Bürger, 1995 a, b; Lande, 1994). Once the deleterious mutation load reaches the point that the best genotype in the population can just replace itself, then any further increase in the mutation load necessarily leads to a reduction in population size. This slightly increases the power of random genetic drift, enhancing the rate of accumulation of subsequent cohorts of mutations, and promoting still further decline in population size. We refer to this synergism, whereby the rate of mutation accumulation increases with the mutation load, as a mutational meltdown.

Although small sexual populations are vulnerable to deleterious-mutation accumulation, the rate of mutational deterioration in the nuclear genome is slowed substantially by segregation and recombination, which enables parents to produce progeny with reduced mutation load at polymorphic loci (Pamilo, Nei & Li, 1995, Charlesworth, Morgan & Charlesworth, 1993; Lande, 1994; Lynch, Conery & Bürger, 1995a, b). The degree to which recombination increases the efficiency of selection appears to be so great that sexual populations with effective sizes in excess of 1000 individuals are essentially invulnerable to a mutational meltdown in the nuclear genome (Schultz & Lynch, 1997).

The possibility remains that even fairly large sexual populations may be vulnerable to deleterious-mutation accumulation in organelles (Hastings, 1992; Gabriel, Lynch & Bürger, 1993). Although organelle genomes can experience recombination (Wolstenholme & Jeon, 1992; Birky, 1995; Gillham, 1995; Lunt & Hyman, 1997), the population-level consequences of such an event are potentially irrelevant. Because organelles are generally inherited uniparentally, recombination among such genomes will usually involve very closely related (most often identical) molecules. However, if a recombination event is to produce an offspring molecule with fitness greater than that of either parent, then each parental molecule must have at least one deleterious mutation at a unique site. Thus, unless heteroplasmic lineages of organelles are preserved for very long periods of time (on the order of thousands of generations), which appears to be ruled out by empirical evidence (Birky, 1995), or unless biparental inheritance commonly combines different organelles into the same individual, recombination among organelles may have

very little impact on the accumulation of deleterious mutations.

Information on the rate and effects of deleterious mutations arising in nuclear genomes derives from laboratory mutation-accumulation experiments in which lines of individuals have been taken through recurrent bottlenecks in order to minimize the effectiveness of selection. Several studies (reviewed in Simmons & Crow, 1977; Lynch & Walsh, 1998) have taken advantage of the lack of recombination in male *Drosophila* as a means for accumulating mutations on autosomes in a clonal fashion. These studies have shown that approximately one new deleterious mutation arises per diploid genome per generation, with the average effect of such mutations on fitness being approximately 2.5% in the heterozygous state. Similar results have recently been obtained with clonally propagated lines of the microcrustacean *Daphnia* (Lynch et al., in prep.) In the plant *Arabidopsis thaliana* (Schultz et al., in prep.) and the nematode *Caenorhabditis elegans* (Keightley & Caballero, 1997, Vassieleva & Lynch, in prep.), the genomic deleterious mutation rate is about an order of magnitude lower, but the average effects are similar.

Because organelles have genome sizes that are on the order of 10^{-5} to 10^{-4} of that of the nuclear genome, mutation-accumulation experiments are unlikely to yield much insight into their mutational properties on reasonable time scales. As an alternative approach to searching for evidence of decreased efficiency of selection against mutations arising in organelle vs. nuclear genomes, we have looked for the signature of such change at the molecular level — an increase in the rate of nucleotide substitution relative to the rate of mutation. The ratio of these two rates, which we refer to as the width of the selective sieve, approaches zero when selection is highly effective at eradicating new mutations and approaches one when the magnitude (or efficiency) of selection is relaxed and genes evolve at the neutral rate. (In principle, the width of the sieve can exceed one when there is positive selection for new mutations.)

As a first test of the idea that the width of the selective sieve is magnified for genes that reside in organelle genomes, comparative studies were performed on the isoaccepting sets of transfer RNA genes that reside in both organelle and nuclear genomes (Lynch, 1996, 1997). These genes provide an ideal substrate for such a test, because their products have similar functions in both the organelle and cytoplasmic environments (primarily translation of messenger RNAs). Moreover, the presence of up to twenty families of tRNA genes

(one for each amino acid) in both genomes of the same individual provides a substantial degree of 'replication' for the analysis. Relative to the situation in the nuclear genome, the width of the selective sieve for tRNA mutations in the mitochondrion is inflated 2.6-fold in mammals and 3.8-fold in invertebrates (Lynch, 1996). In fungi and plants, the width of the sieve is inflated 2.5-fold and 13-fold, respectively, in the mitochondrion, and there is a 1.3-fold inflation in the plant chloroplast (Lynch, 1997).

Thus, for tRNA genes, results from a wide phylogenetic array of species consistently indicate that the efficiency of selection against mitochondrial mutations is reduced relative to that which occurs in the nuclear genome. Results cited in Lynch (1996, 1997) support the idea that some of the excess mutations that are accumulating in organelle tRNAs are indeed deleterious. Nuclear tRNAs are among the most evolutionarily stable of all known genes, with 13 of the approximately 70 nucleotide sites (exclusive of the anticodon sites) being invariant across all eukaryotes and prokaryotes and across all families of tRNAs (Söll & RajBhandary, 1995). This implies that natural selection plays a very important role in maintaining the optimal molecular architecture of these genes. Yet, no such invariant sites are found within mitochondrial tRNAs. Relative to their nuclear counterparts, organelle tRNAs also exhibit a substantial reduction of stem duplex stabilities, due largely to mutation pressure towards A-U vs. G-C bonds, and the loops of organelle tRNAs are exceptionally variable in size.

If the hypothesis that organelle genomes are gradually accumulating mildly deleterious mutations as a consequence of repressed recombination is correct, then the patterns observed for tRNAs should extend to other classes of genes. Thus, the purpose of this paper is to expand our previous study to protein-coding loci and to the ribosomal RNAs. Our results show that the width of the selective sieve in organelle genomes is magnified on a genome-wide basis, and suggest that this is largely a consequence of a reduction in the effective population size for organelle genes.

Methods

The following analyses are based on DNA sequence comparisons across major phylogenetic groups. Mammalian analyses involved the three pairwise comparisons among artiodactyls, primates, and rodents. Anamniote analyses (confined to proteins, due to the

absence of complete rRNA sequences) compared ray-finned fishes and amphibia, while invertebrate analyses compared arthropods and nematodes. Plant analyses compared monocots and dicots, while fungal analyses were restricted to *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*.

Source of data

The sequences for all of the protein coding genes in this study were accessed from the Genbank data repository. For each of the five major groups noted above, all available sequences were downloaded, and alignments of the amino acid sequences were accomplished with the assistance of ClustalW (Thompson, Higgins & Gibson, 1994), occasionally with minor subsequent adjustments made by eye. These were then converted to nucleotide sequence alignments, again with occasional minor adjustments made by eye. Only the portions of the sequences that were unambiguously aligned were retained for analysis.

Small and large subunit ribosomal sequences were retrieved from the rRNA WWW server at URL <http://rrna.uia.ac.be/> (release - Jan. 14, 1997). This database contains rRNA secondary structure information embedded into aligned rRNA sequences (De Rijk, Van de Peer & Wachter, 1997; Van de Peer et al., 1997). To confine comparisons between homologous sites, we used the tkDCSE routine (De Rijk & De Wachter, 1993) to compile files consisting only of stem regions. Pseudoknots within stems are also included in our analysis. A couple of rRNA sequences not found on the rRNA WWW server were added to the alignment using ClustalV (Higgins & Sharp, 1989). In a few instances, blocks of the alignment supplied by the rRNA WWW server were realigned within individual data files by ClustalV.

Rates of nucleotide substitution

As estimates of the numbers of mutations per nucleotide site, we employed estimates of the numbers of synonymous substitutions observed at four-fold degenerate sites in protein-coding genes, i.e., at codon pairs that were identical between taxa in the first two positions and for which the nucleotide in the third position does not affect the reading of the genetic code. To account for multiple substitutions per nucleotide site, a modification of the equation of Tajima (1993), given as the series-expansion Equation (1) of Lynch (1996), was employed as an estimator of the mean number

of substitutions per site. This formula reduces bias in estimates of genetic distance that can result with expressions that ignore sampling error, and the general approach we have taken yields essentially unbiased results (Comeron, 1995). For each phylogenetic comparison, the number of synonymous substitutions per synonymous site was estimated for each particular gene by averaging the results over all internal pairwise comparisons. For example, if there were three fish and four amphibian sequences, the twelve pairwise estimates would be averaged to give the overall estimate for anamniotes. The final estimates of the synonymous substitution rates were then obtained by averaging the results over all of the protein-coding loci that were analyzed.

For the animal, fungal, and plant mitochondria, every protein-coding gene in the genome was employed in this study, whereas results for the chloroplast were obtained by using a sample of 25 of the approximately 100 protein-coding genes in this genome. Mutation (synonymous substitution) rate estimates for the anamniote, invertebrate, fungal, and plant nuclear genomes were obtained as averages of 25 hapazardly procured genes from Genbank (available from the authors on request). Synonymous substitution rates do not vary greatly from gene to gene, and for no phylogenetic group does the standard error of our mutation rate estimate exceed 8% of the estimated value. For the rate of synonymous substitution in the mammalian nuclear genome, the average results given in Li and Graur (1991), Eastal and Collet (1994), and Ohta (1995) were used; collectively these studies have surveyed several dozens of loci, and all three studies have obtained very similar results.

To estimate the number of nonsynonymous substitutions per nonsynonymous site in protein-coding loci, we used the method of Nei and Gojobori (1986) as implemented in MEGA (Kumar et al., 1993). As with our estimator of the number of synonymous substitutions per synonymous site, this method accounts for multiple substitutions per site. To estimate the number of substitutions per site in ribosomal DNA sequences, we again used Equation (1) in Lynch (1996).

Width of the selective sieve

Since the rate of molecular evolution is a function of both the mutation rate and the efficiency of selection, interpretations of observed rates of substitution need to consider the rate expected in the absence of selection. For protein-coding genes, the usual approach is

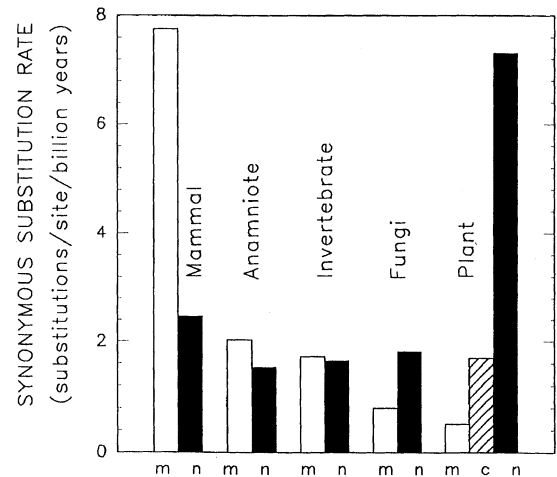


Figure 1. Estimates of the rate of synonymous substitution at four-fold degenerate sites in protein-coding loci. Estimates of numbers of substitutions per site were converted to rates by using the divergence times given in Lynch (1997), except in the case of fungi, which are simply reported as observed numbers per site (after correcting for multiple substitutions per site). Genomes are denoted as: mitochondrial (m), chloroplast (c), and nuclear (n). The standard errors of all estimates are less than 8% of the plotted values.

to compare the rates of synonymous and nonsynonymous substitution, the former serving as an estimate of the neutral rate of molecular evolution (which ideally is equivalent to the mutation rate for nucleotide substitutions, as assumed herein). We adhere to this approach with protein-coding genes. To extend this idea to tRNA and rRNA genes, we employ the ratio of the observed substitution rate in such genes to the rate of synonymous substitution in protein-coding genes. In the following, we refer to all such ratios as widths of the selective sieve. To obtain standard errors of such estimates, we employed the Taylor-expansion expression for the sampling variance of a ratio (Appendix 1 in Lynch & Walsh, 1998).

Results

Substantial variation exists among genomes in the mutation rate as estimated from numbers of synonymous substitutions per synonymous site (Figure 1). The inflation of the mutation rate in organelle genomes, commonly alluded to in the literature, is primarily a vertebrate phenomenon. The difference in mitochondrial and nuclear rates is nonsignificant in invertebrates, while in fungi there is a two-fold decrease in the rate in the mitochondrial genome, and in plants the ratio

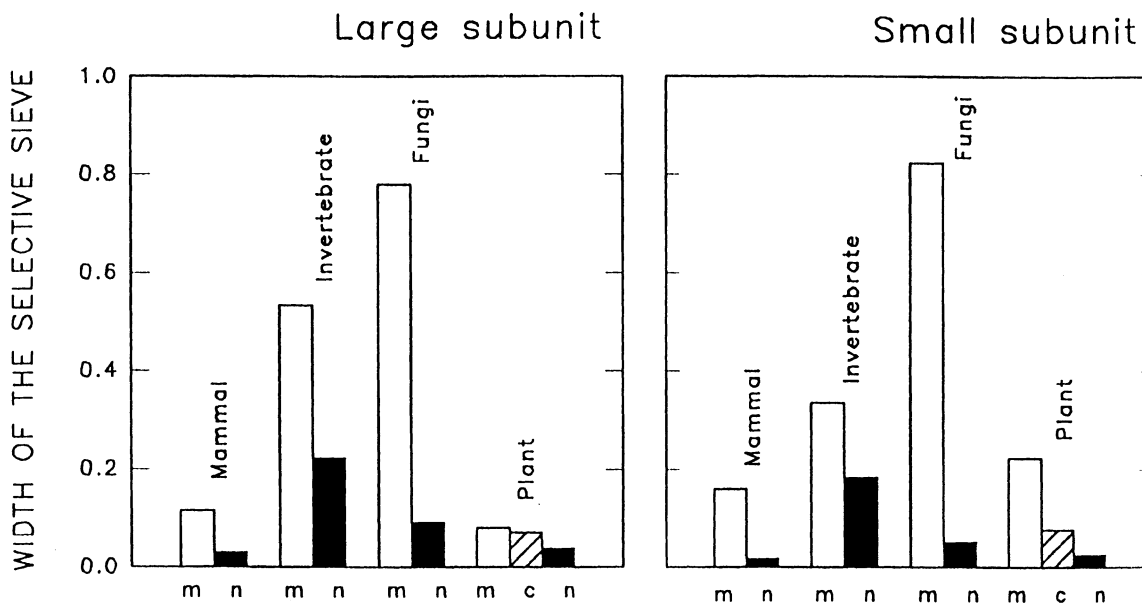


Figure 2. Average width of the selective sieve for the small and large subunits of the ribosomal RNA genes. All differences within taxonomic groups are significant at the 0.01 level.

of rates is approximately 1 : 3 : 14 for mitochondrion : chloroplast : nucleus. With such variation in the synonymous substitution rate, it is clear that if any inferences about the efficiency of selection are to be drawn from observed rates of nucleotide substitutions, they must take into consideration the differences among taxa in the neutral expectation.

For both the large and small subunit ribosomal RNAs, there is an inflation in the width of the selective sieve in the organelle vs. nuclear genomes in every phylogenetic group included in this study (Figure 2). For the mitochondrial genome, the inflation averages approximately six-fold in mammals, two-fold in invertebrates, twelve-fold in fungi, and five-fold in plants, while for the plant chloroplast there is an approximately two-fold inflation.

For protein-coding genes, relative to the situation in the nuclear genome, there is an approximately four-fold inflation in the width of the selective sieve in the mitochondrial genome of plants and an approximately 50% inflation in the chloroplast (Figure 3). For the mitochondrial genomes of invertebrates and fungi, there is an approximately two-fold inflation. The data suggest an approximately 25% reduction in the width of the sieve in vertebrate mitochondrial genomes, but as noted in the discussion, the vertebrate nuclear results are likely to be upwardly biased.

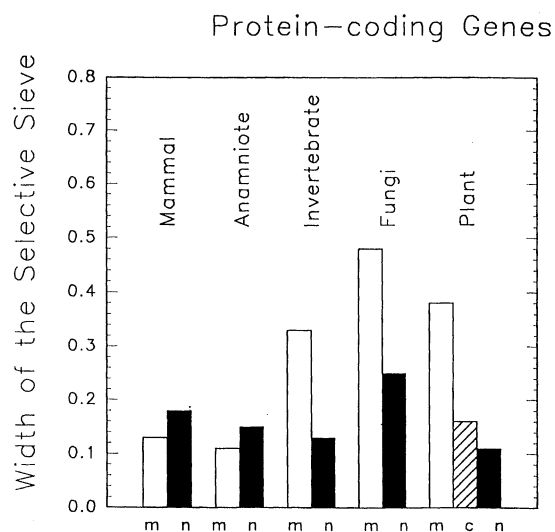


Figure 3. Average width of the selective sieve for protein-coding genes. All differences within taxonomic groups are significant at the 0.01 level.

Discussion

The rate of long-term molecular evolution is equal to the product of the number of mutations arising per site per population and the probability of fixation of a new mutation. The closest that we can come to an estimate of the mutation rate for protein-coding genes

is the substitution rate at four-fold degenerate sites. In principle, such rates may be underestimates of the true mutation rate if there is codon selection for particular third-position nucleotides. In the following discussion, we refer to our estimates of the synonymous (silent) substitution rate as measures of the mutation rate with this qualification in mind.

As can be seen in Figure 1, a substantial amount of variation in the evolutionary rate among genomes is a consequence of differences in rates of silent substitution. As a ratio of the observed rate of nucleotide substitution and the neutral expectation, the width of the selective sieve separates out the effects of selection from the effects of mutation on the evolutionary rate, yielding an estimate of the ratio of fixation probabilities of observed and neutral substitutions. More precisely, the width of the sieve is equivalent to the ratio of fixation probabilities for all mutations and that of synonymous substitutions. Our application of the width of the selective sieve to tRNA and rRNA genes is a straightforward extension of the ratio of nonsynonymous to synonymous substitution rates commonly applied in studies of protein evolution (Li & Graur, 1991).

The results herein, combined with those of Lynch (1996, 1997), provide compelling evidence that newly arisen mutations in most organelle genes have a higher probability of fixation than do those that arise in nuclear genes. This general principle applies to a very wide array of phylogenetic groups for gene families as disparate as transfer RNAs, ribosomal RNAs, and a wide array of protein-coding loci. Averaging over our results for these three gene types, the widths of the selective sieve for mitochondrial and nuclear genomes are 0.15 and 0.09 in mammals, 0.38 and 0.14 in invertebrates, 0.59 and 0.17 in fungi, and 0.57 and 0.08 in plants, while the average width for the chloroplast is 0.12. Taking all of these results into consideration, it appears that mutations that modify the products of genes residing in an organelle experience, on average, a two- to seven-fold increase in the fixation probability relative to that for nuclear mutations. If selectively driven codon bias is greater in nuclear genes, which seems likely given the arguments below, then the true situation is even more extreme.

Although our observations for protein-coding genes in vertebrates appear to be an exception to this pattern, the results for protein-coding loci are expected to be less reliable than those for tRNAs and rRNAs. In the latter two cases, the units of comparison among genomes have similar functions in both genetic envi-

ronments, but because the same protein-coding genes are not found in both organelle and nuclear genomes, we were forced to make a more or less haphazard selection of nuclear gene sequences from those available in the existing database. Some of the nuclear genes that we employed were highly conserved histones and cytoskeletal proteins, which may have biased the average widths of the selective nuclear sieve in a downward fashion. However, for vertebrates, a substantial amount of upward bias may have occurred. Prior to the emergence of the major vertebrate lineages (around 500 million years ago), there were two complete genome duplications (Holland et al., 1994). A large fraction of the duplicate genes created by these events have not been silenced, and, as a consequence, many pairs of genes that have been sequenced and regarded as homologous are probably ancient paralogues. Paralogous genes may be somewhat released from selection as a consequence of their redundancy or they may be selected for new gene functions. Either situation would inflate the rate of nonsynonymous substitution relative to the expectation for a single-copy gene. Thus, we do not regard our results for vertebrate protein-coding genes as providing a convincing contradiction to the general rule that the width of the selective sieve is inflated in nonrecombining organelles.

Although our analyses are perhaps the most comprehensive in terms of phylogenetic and gene family coverage, numerous other lines of evidence support the idea that nonrecombining genomes accumulate deleterious mutations, particularly when exposed to periodic bottlenecks in population size. First, such results have been observed directly in short-term laboratory experiments with RNA viruses (Chao, 1990; Duarte et al., 1994; Escarmis et al., 1996), the bacterium *Escherichia coli* (Kibota & Lynch, 1996), and an artificially constructed nonrecombining drosophilid chromosome (Rice, 1994). Second, population surveys of mammalian and *Drosophila* mitochondrial genomes have consistently revealed an excess of nonsynonymous polymorphisms within species relative to expectations based on between-species comparisons (Ballard & Kreitman, 1994; Nachman Boyer & Aquadro, 1994; Nachman et al., 1996; Rand, Dorfsman & Kann, 1994; Rand & Kann, 1996). The favored interpretation of this observation is that recurrent mutation creates deleterious alleles with effects that are mild enough to allow drift to observable frequencies but strong enough to prevent fixation. However, it seems likely that deleterious mutations that are capable of rising to observable frequencies must also be subject to fixation. Third,

comparative surveys indicate that the genes contained within endosymbiotic bacteria in insects exhibit a substantial increase in the width of the selective sieve relative to that in free-living bacteria (Moran, 1996). Like organelles, these endosymbiotic bacteria are inherited uniparentally, with effectively zero recombination and with substantial population bottlenecks occurring during transmission to host progeny. Fourth, a substantial body of data points to a causal connection between mitochondrial mutations and a diverse array of human genetic disorders (Wallace, 1992, 1994). Mildly deleterious mitochondrial mutations have been observed to have frequencies as high as 5%.

There are at least three possible reasons for the enhanced fixation probabilities for deleterious mutations in organelle-encoded genes. First, there is the basic premise that the efficiency of selection is reduced by the inevitable build-up of linkage disequilibrium that arises in nonrecombining genomes. Second, because they are haploid and uniparentally inherited, organelle genes have effective population sizes that are approximately one-quarter of those for nuclear genes in the same species. The resultant increase in random genetic drift for organelle genes will necessarily inflate the rate of fixation of mildly deleterious mutations to some degree. Third, it has been argued that a reduction in the rate of elimination of mutant alleles in the mitochondrion might be due, in part, to a relaxation of functional constraints on mitochondrial genes relative to their nuclear counterparts (Brown et al., 1982; Kumazawa & Nishida, 1993).

A rough evaluation of the hypothesis of relaxed selection can be made by considering the standard diffusion approximation for the fixation of a mildly deleterious mutation (Crow & Kimura, 1970), which assumes additive gene effects. For an ideal random-mating population with effective size N , the width of the selective sieve (the ratio of the substitution rate relative to the neutral expectation) for a deleterious nuclear mutation is approximately

$$\omega_n = \frac{2s_n N}{e^{2s_n N} - 1}, \quad (1 \text{ a})$$

where s_n is the fractional reduction of fitness in a homozygous mutant. On the other hand, for a population with a 1:1 sex ratio, the width of the selective sieve for a uniparentally inherited organelle genome is approximately

$$\omega_o = \frac{s_o N}{e^{s_o N} - 1}, \quad (1 \text{ b})$$

where N again denotes the effective population size for the nuclear genome, and s_o denotes the reduction of fitness in a homoplasmic mutant. By substituting estimates of the width of the selective sieve for ω_o and ω_n , estimates of the composite quantities $s_o N$ and $s_n N$ can be obtained. Taking the ratio of these estimates for organelle and nuclear genes factors out population size and provides a measure of the average selection coefficient against an organelle mutation relative to that against a nuclear mutation.

Such estimates of s_o/s_n should only be viewed as rough approximations for the following reasons. First, because the preceding expressions ignore complications due to linkage disequilibrium, they will yield downwardly biased estimates of sN for a given ω (Lynch, Conery & Bürger, 1995a, b), but the bias is expected to be more pronounced for genes in organelles. Second, failure to account for nonadditivity of mutational effects will also downwardly bias the estimates of sN , but again to a degree that is likely to be greater for organelle genes. Mildly deleterious mutations are usually slightly recessive (Lynch & Walsh, 1997), and this enhances their probability of fixation by sheltering them from selection when at low frequency. Because the number of organelle genomes per cell is typically on the order of tens to thousands (compared to two for diploid nuclear genomes), the phenotypic effects of a single mutation may be essentially zero until the intragenomic frequency has risen to a moderate level (Shoubridge, 1994). On the other hand, Equation (1b) does not explicitly account for the selection that may occur on an organelle mutation in the heteroplasmic state, i.e., prior to fixation within its organelle lineage (Takahata & Slatkin, 1983), and hence may overestimate $s_o N$ somewhat.

Application of the average estimates of ω for tRNA, rRNA, and protein-coding genes (the data in Figures 2, 3; and in Lynch, 1996, 1997) yields estimates of sN for nuclear-encoded genes in the range of 1.2 to 2.7 for all of the phylogenetic groups employed in this study. The range for organelle-encoded genes is similar, 1.3 to 3.4, with two exceptions. First, as noted in Lynch (1997), the width of the selective sieve for plant mitochondrial tRNAs is not significantly different from one, so $sN \simeq 0$. Second, the exceptionally high width of the selective sieve for ribosomal RNA subunits in fungal mitochondria implies that $sN \simeq 0.2$ for these genes.

The estimated ratios s_o/s_n all exceed 0.9, although not greatly so, except in the case of plant mitochondrial tRNAs and fungal mitochondrial rRNAs where they

Table 1. Estimates of the relative selection coefficients against organelle and nuclear mutations and of the ratios of observed widths of the selective sieve to expectations based on observations for nuclear genes

	Transfer RNAs	Ribosomal RNAs	Protein- coding genes
Relative selection coefficients (s_o/s_n)			
Mammal mt	1.40	1.19	2.29
Anamniote mt	–	–	2.27
Invertebrate mt	0.94	1.15	1.18
Fungal mt	1.00	0.20	1.08
Plant mt	0.00	1.24	1.00
Plant cp	1.78	1.60	1.76
Ratio of observed to predicted widths of the selective sieve for organelle mutations			
Mammal mt	0.58	0.71	0.28
Anamniote mt	–	–	0.26
Invertebrate mt	1.09	0.89	0.81
Fungal mt	1.03	2.56	0.93
Plant mt	3.53	0.68	1.00
Plant cp	1.03	0.99	1.02

are 0.0 and 0.2, respectively (upper half of Table 1). Although these computations are crude, they are quite consistent in suggesting that the strength of selection on organelle mutations is of the same order of magnitude of that operating on nuclear-gene mutations. Thus, there is no compelling evidence that the enhanced rate of evolution of organelle genomes is a product of relaxed selection against new mutations, except perhaps in the case of plant mitochondrial tRNAs and fungal mitochondrial rRNAs. This suggests that a reduction in the *efficiency* of selection, either due to a reduction in effective population size or an increase in linkage disequilibrium, or both, is involved.

Further application of Equations 1a, b can yield some qualitative insight into the mechanisms that reduce the efficiency of selection in organelle genomes. Although Equation 1b accounts for the reduction in the effective population size for an organelle genome resulting from uniparental inheritance, it does not take into consideration the effects of linkage disequilibrium. Therefore, if we suppose that the estimates of $s_n N$ obtained for nuclear genomes are reasonably representative of the true values of $s_o N$ for organelle genomes in the same phylogenetic group, then substitution of estimates of $s_n N$ for $s_o N$ in Equation 1b should provide an estimate of the width of the selective sieve for the organelle genome under the assumption that the

reduction in the efficiency of selection is solely a consequence of a reduction in effective population size.

The actual values of ω_o are fairly close to these expectations, with a few exceptions (lower half of Table 1). For the plant mitochondrial tRNAs and fungal mitochondrial rRNAs, the observed ω_o are in substantial excess of the predictions based on the nuclear results. In principle, such inflation could be a consequence of the negative effects of organelle genome-wide linkage disequilibrium on the efficiency of selection, but this seems to be ruled out by the fact that other classes of genes in plant and fungal organelles have estimates of ω_o that are very close to expectations based solely on a reduction in effective population size (Table 1). Thus, it seems likely that purifying selection on these genes is relaxed, i.e., $s_o \leq s_n$. On the other hand, for the vertebrate protein-coding genes, the observed ω_o are about four times lower than the predicted value, but bias in this direction is expected on the basis of the arguments presented above. If these exceptions are ignored, the ratios of observed to expected ω average (\pm SE) to 0.95 ± 0.04 . Thus, the data suggest that most of the reduction in the efficiency of selection against mildly deleterious mutations in organelle genomes is a simple consequence of a reduction in population size.

With these results in hand, it is possible to make some rough calculations of the long-term consequences of deleterious mutations in organelles. Using the diffusion approximation for the fixation probability of a deleterious mutation in an organelle, $2s/(e^{sN} - 1)$, and noting that the number of new mutations arising in a population per year is $UN/2$, where s is the selection coefficient, U is the genomic mutation rate per organelle per year, and N is the total effective population size (males and females), then the fractional reduction in mean population fitness per year is

$$\Delta \bar{W} = \frac{UNs^2}{e^{sN} - 1}. \quad (2)$$

(If there are n organelle genomes per zygote, and each new mutation within an individual has probability $1/n$ of drifting to fixation, then the number of newly arising homoplasmic lineages of mutations is $(UNn/2) \cdot (1/n) = UN/2$ per generation; so multiplicity of organelle genomes has no influence on this result so long as the force of intragenomic selection is weak relative to that operating at the intergenomic level.)

Recalling the rates of synonymous substitution per nucleotide site given in Figure 1 and noting that organelle genomes contain on the order of 10^4 (animal

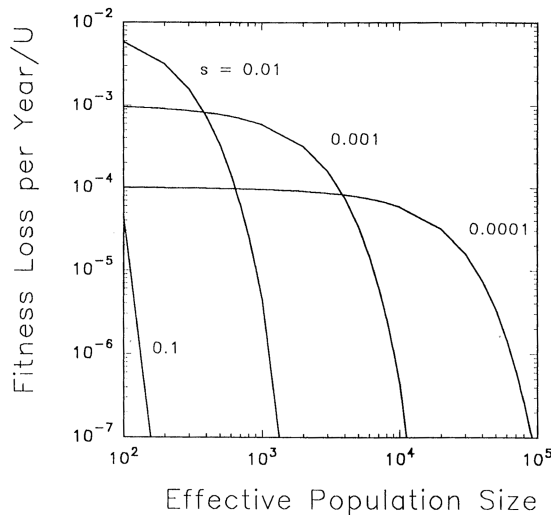


Figure 4. Expected fractional loss of fitness per year (scaled by the organelle genomic mutation rate) as a function of the selection coefficient (s) and the effective population size (N).

mitochondria) to 10^5 (plant organelles) nucleotides, then U is on the order of 10^{-4} per year. Our preceding analyses suggest that the average value of s for an organelle mutation is similar to that for nuclear mutations, in which case (to an order of magnitude) \bar{s} is likely to be on the order of 0.01 (Simmons & Crow, 1977; Lynch & Walsh, 1998). The data are not extensive, but it seems likely that the actual distribution of s is highly L-shaped, with mutations with very minor effects having the highest density (Keightley, 1994). The results in Figure 4 show how the fractional loss of fitness scaled by the mutation pressure, $\Delta\bar{W}/U = Ns^2/(e^{sN} - 1)$, depends on the selection pressure (s) and the effective population size (N) for likely values of s . To know the total effects of mutations on population fitness, one would need to integrate these curves over the actual distribution of s . Thus, depending on the exact form of the distribution of s , these results suggest that $\Delta\bar{W}/U$ is on the order of 10^{-3} per year or smaller, declining with increasing population size (Figure 4). Consider, for example, a species with a long-term effective size of $N = 10^4$. Assuming a high density of mutational effects (s_0) in the range of 10^{-4} to 10^{-3} , then the fractional loss of fitness per generation is on the order of $10^{-6}U$ to $10^{-4}U$ per year, which translates into an approximate fitness loss of 0.01% to 1% per million years when $U = 10^{-4}$.

Thus, it does not seem likely that deleterious mutation accumulation in organelles imperils species on time scales less than one million years, provided the

long-term effective population size exceeds several thousands of individuals. If the reduction in fitness is only 0.01% per million years, then even after a billion years of mutation accumulation (the approximate time since the divergence of animals, fungi, and plants), only a 10% loss of fitness would have occurred. If, on the other hand, the rate of loss is on the order of 1% per million years, then the expected loss of fitness after 10 million years (the average of lifespan of a species) is approximately 10%, after 100 million years (the approximate time since the origin of the orders of mammals) is approximately 73%, and after 0.5 billion years (the approximate time since the origin of the major animal phyla) is 99.3%.

These computations do not take into account the substantial opportunities for compensatory mutations and/or back mutations that exist on such large time scales. However, our approximate results also only refer to the expected loss of fitness. Due to the stochastic nature of the mutational process, some species will decline at a much slower rate and others at a much faster rate. Moreover, it should be noted that our estimated genomic mutation rates and average effects of mutations in organelles are based on lineages of organisms that have survived over very long periods of time. Thus, it is plausible that deleterious mutation accumulation in organelle genomes, combined with differential extinction and proliferation of species lineages, has played some role in the patterning of the deep phylogenetic organization of life.

Acknowledgements

We thank C. Aquadro for helpful comments. The work was supported by NIH grant RO1-GM36827 to ML.

References

- Ballard, J.W.O. & M. Kreitman, 1994. Unraveling selection in the mitochondrial genome of *Drosophila*. *Genetics* 138: 757–772.
- Birky, C.W., Jr., 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proc. Natl. Acad. Sci. USA* 92: 11331–11338.
- Birky, C.W., Jr. & J.B. Walsh, 1988. Effects of linkage on rates of molecular evolution. *Proc. Natl. Acad. Sci. USA* 85: 6414–6418.
- Brown, W.M., E.M. Prager, A. Wang & A.C. Wilson, 1982. Mitochondrial DNA sequences of primates: tempo and mode of evolution. *J. Mol. Evol.* 18: 225–239.
- Chao, L., 1990. Fitness of RNA virus decreased by Muller's ratchet. *Nature* 348: 454–455.

- Charlesworth, D., M.T. Morgan & B. Charlesworth., 1993. Mutation accumulation in finite outbreeding and inbreeding populations. *Genet. Res.* 61: 39–56.
- Comeron, J.M., 1995. A method for estimating the numbers of synonymous and nonsynonymous substitutions per site. *J. Mol. Evol.* 41: 1152–1159.
- Crow, J.F. & M. Kimura, 1970. *An Introduction to Population Genetics Theory*. Harper and Row, New York.
- De Rijk, P., Y. Van de Peer & R. De Wachter, 1997. Database on the structure of large ribosomal subunit RNA. *Nucl. Acids Res.* 25: 117–122.
- De Rijk, P. & R. De Wachter, 1993. DCSE, an interactive tool for sequence alignment and secondary structure research. *Comput. Appl. Biosci.* 9: 735–740.
- Duarte, E.A., I.S. Novella, S. Ledesma, D.K. Clarke, A. Moya, S.F. Elena, E. Domingo & J.J. Holland, 1994. Subclonal components of consensus fitness in an RNA virus clone. *J. Virol.* 68: 4295–4301.
- Eastale, S. & C. Collet, 1994. Consistent variation in amino-acid substitution rate, despite uniformity of mutation rate: protein evolution in mammals is not neutral. *Mol. Biol. Evol.* 11: 643–647.
- Escarmis, C.M. Davila, N. Charpentier, A. Bracho, A. Moya & E. Domingo, 1996. Genetic lesions associated with Muller's ratchet in an RNA virus. *J. Mol. Biol.* 264: 255–267.
- Felsenstein, J., 1974. The evolutionary advantage of recombination. *Genetics* 78: 737–756.
- Gabriel, W., M. Lynch & R. Bürger, 1993. Muller's ratchet and mutational meltdowns. *Evolution* 47: 1744–1757.
- Gessler, D.D.G., 1996. The constraints of finite size in asexual populations and the rate of the ratchet. *Genet. Res.* 66: 241–253.
- Gillham, N.W., 1995. *Organelle Genes and Genomes*. Oxford Univ. Press, New York.
- Haigh, J., 1978. The accumulation of deleterious genes in a population. *Theor. Pop. Biol.* 14: 251–267.
- Hastings, I.M., 1992. Population genetic aspects of deleterious cytoplasmic genomes and their effect on the evolution of sexual reproduction. *Genet. Res.* 59: 215–225.
- Higgins, D.E. & P.M. Sharp, 1989. Fast and sensitive multiple sequence alignments on a microcomputer. *CABIOS* 5: 151–153.
- Higgs, P.G., 1994. Error thresholds and stationary mutant distributions in multi-locus diploid genetics models. *Genet. Res.* 63: 63–78.
- Hill, W.G. & A. Robertson, 1966. The effect of linkage on limits to artificial selection. *Genet. Res.* 8: 269–294.
- Holland, P.W.H., J. Garcia-Fernandez, N.A. Williams & A. Sidow, 1994. Gene duplications and the origins of vertebrate development. *Devel. Suppl.* 125–33.
- Keightley, P.D., 1994. The distribution of mutation effects on viability in *D. melanogaster*. *Genetics* 138: 1315–1322.
- Keightley, P.D. & A. Caballero, 1997. Genomic mutation rates for lifetime reproductive output and lifespan in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* 94: 3823–3827.
- Kibota, T.T. & M. Lynch, 1996. Estimate of the genomic mutation rate deleterious to overall fitness in *Escherichia coli*. *Nature* 381: 694–696.
- Kumar, S., K. Tamura & M. Nei, 1993. MEGA: Molecular Evolutionary Genetics Analysis, Version 1.01. Penn. State Univ., University Park, PA.
- Kumazawa, Y. & M. Nishida, 1993. Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. *J. Mol. Evol.* 37: 380–398.
- Lande, R., 1994. Risk of population extinction from new deleterious mutations. *Evolution* 48: 1460–1469.
- Li, W.-H. & D. Graur, 1991. *Fundamentals of Molecular Evolution*. Sinauer Assocs., Sunderland, MA.
- Lunt, D.H. & B.C. Hyman, 1997. Animal mitochondrial DNA recombination. *Nature* 387: 247.
- Lynch, M., 1996. Mutation accumulation in transfer RNAs: molecular evidence for Muller's ratchet in mitochondrial genomes. *Mol. Biol. Evol.* 13: 209–220.
- Lynch, M., 1997. Mutation accumulation in nuclear, organelle, and prokaryotic transfer RNA genes. *Mol. Biol. Evol.* 14: 914–925.
- Lynch, M., R. Bürger, D. Butcher & W. Gabriel, 1993. The mutational meltdown in asexual populations. *J. Heredity* 84: 339–344.
- Lynch, M., J. Conery & R. Bürger, 1995a. Mutational meltdowns in sexual populations. *Evolution* 49: 1067–1080.
- Lynch, M., J. Conery & R. Bürger, 1995b. Mutation accumulation and the extinction of small populations. *Amer. Nat.* 146: 489–518.
- Lynch, M. & W. Gabriel, 1990. Mutation load and the survival of small populations. *Evolution* 44: 1725–1737.
- Lynch, M. & J.B. Walsh, 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Assocs., Inc., Sunderland, MA.
- Moran, N.A., 1996. Accelerated evolution and Muller's ratchet in endosymbiotic bacteria. *Proc. Natl. Acad. Sci. USA* 96: 2873–2878.
- Muller, H.J., 1964. The relation of recombination to mutational advance. *Mut. Res.* 1: 2–9.
- Nachman, M.W., S.N. Boyer & C.F. Aquadro, 1994. Non-neutral evolution at the mitochondrial ND3 gene in mice. *Proc. Natl. Acad. Sci. USA* 91: 6364–6368.
- Nachman, M.W., W.M. Brown, M. Stoneking & C.F. Aquadro, 1996. Nonneutral mitochondrial DNA variation in humans and chimpanzees. *Genetics* 142: 953–963.
- Nei, M. & T. Gojobori, 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* 3: 418–426.
- Ohta, T., 1995. Synonymous and nonsynonymous substitutions in mammalian genes and the nearly neutral theory. *J. Mol. Evol.* 40: 56–63.
- Pamilo, P., M. Nei & W.-H. Li, 1987. Accumulation of mutations in sexual and asexual populations. *Genet. Res.* 49: 135–146.
- Rand, D.M., M. Dorfsman & L.M. Kann, 1994. Neutral and non-neutral evolution of *Drosophila* mitochondrial DNA. *Genetics* 138: 741–756.
- Rand, D.M. & L.M. Kann, 1996. Excess amino acid polymorphism in mitochondrial DNA: contrasts among genes from *Drosophila*, mice, and humans. *Mol. Biol. Evol.* 13: 735–748.
- Rice, W.R., 1994. Degeneration of a nonrecombining chromosome. *Science* 263: 230–232.
- Schultz, S.T. & M. Lynch, 1997. Deleterious mutation and extinction: effects of variable mutational effects, synergistic epistasis, beneficial mutations, and degree of outcrossing. *Evolution* 51: 1363–1371.
- Shoubridge, E.A., 1994. Mitochondrial DNA diseases: histological and cellular studies. *J. Bioenerg. Biomem.* 26: 301–310.
- Simmons, M.J. & J.F. Crow, 1977. Mutations affecting fitness in *Drosophila* populations. *Ann. Rev. Genet.* 11: 49–78.
- Söll, D. & U.L. RajBhandary, 1995. *tRNA: structure, biosynthesis, and function*. ASM Press, Washington, D.C.
- Stephan, W., L. Chao & J.G. Smale, 1993. The advance of Muller's ratchet in a haploid asexual population: approximate solutions based on diffusion theory. *Genet. Res.* 61: 225–231.
- Tajima, F., 1993. Unbiased estimation of evolutionary distance between nucleotide sequences. *Mol. Biol. Evol.* 10: 677–688.
- Takahata, N. & M. Slatkin, 1983. Evolutionary dynamics of extranuclear genes. *Genet. Res.* 42: 257–265.

- Thompson, J.D., D.G. Higgins & T.J. Gibson, 1994. Clustal-W: improving the sensitivity of practical multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* 22: 4673–4680.
- Van de Peer, Y., J. Jansen, P. De Rijk & R. De Wachter, 1997. Database on the structure of small ribosomal subunit RNA. *Nucl. Acids Res.* 25: 111–116.
- Wallace, D.C., 1992. Diseases of the mitochondrial DNA. *Annu. Rev. Biochem.* 61: 1175–1212.
- Wallace, D.C., 1994. Mitochondrial DNA mutations in diseases of energy metabolism. *J. Bioenerg. Biomem.* 26: 241–1250.
- Wolstenholme, D.R. & K.W. Jeon (eds.), 1992. *Mitochondrial Genomes*. Academic Press, New York.