Genetica **102/103:** 71–81, 1998.

Requisite mutational load, pathway epistasis, and deterministic mutation accumulation in sexual versus asexual populations

William R. Rice
Department of Biology, University of California, Earth and Marine Science Building, 1156 High Street, Santa Cruz, CA 95064, USA

Key words: mutational load, mutation accumulation, epistasis, recombination

Abstract

A measure of the equilibrium load of deleterious mutations is developed that explicitly incorporates the level of genome-wide linkage disequilibrium. This measure, called the requisite mutational load, is based on the minimal net reproductive rate of the least mutated class necessary to prevent deterministic mutation accumulation. If this minimal net reproductive rate is larger than ecological or physiological constraints allow, then: a) the population is driven to extinction via deterministic mutation accumulation, or b) a mutational Red-Queen ensues with adaptation counterbalancing mutation accumulation. Two population parameters determine the requisite mutational load: a) the equilibrium strength of selection, measured as a selection gradient, and b) the equilibrium opportunity for selection, measured as the variance in number of mutations per genome. The opportunity for selection is decomposed into the accumulation of mutations (average number per genome) and the level of genome-wide linkage disequilibrium. Recombination can substantially reduce the requisite mutational load, compared to clonal reproduction, when there is buffering and/or reinforcing epistasis and also when there is positive assortative mating for fitness. Recombination is advantageous because it reduces the negative (variance reducing) linkage disequilibrium induced by beneficial epistasis. The functional form of the expression for requisite mutational load illustrates why epistasis within pathways, i.e., among closely interacting genes, is a powerful alternative to genome-wide truncation selection, as a means of reducing mutational load.

Introduction

The concept of mutational load ($L = [\overline{W}_{unmutated} - \overline{W}_{realized}] \; / \; \overline{W}_{unmutated}$) has been used to motivate a major advantage to sexual recombination (Kimura & Maruyama, 1966; Crow & Simmons, 1983; Kondroshov, 1988). The biological significance of mutational load, however, has been disputed for several reasons. First, L is calibrated relative to an undefined and unmeasurable standard, i.e., to a hypothetical, mutation-free genotype. Second, compensatory factors (e.g., increased survival of sibs when competition for parental investment is reduced by the death of littermates) may make the demographic impact of mutational load far smaller than predicted by theory. Third, theory concerning reduced mutational load in recombining populations is based predominantly on genome-wide truncation selection. This emphasis on truncation selection seems to be inspired more by its theoretical utility in reducing mutational load than any substantive ecological evidence for its operation in nature.

Here a new measure of mutational load is developed. It is motivated by a fundamental difference (genetic polarization, see below) in the structure of asexual vs. sexual populations. The new measure is based on the requisite net reproductive rate of the most fit mutational class (that is actually present in a population), which is needed to prevent deterministic mutation accumulation. This measure is used to suggest how epistasis within pathways (enzymatic, signal transduction, developmental, and so forth) can provide a strong selective advantage to sexual recombination.

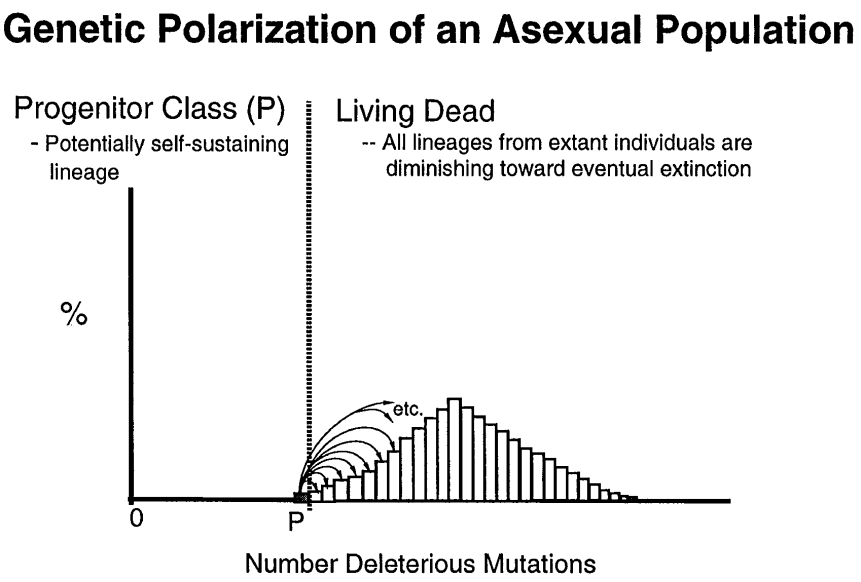

Figure 1. Genetic polarization dichotomizes the population into a self-sustaining progenitor class (*P*) and the non-self-sustaining living dead. The distribution shown is arbitrary, because the actual shape depends on the form and extent of epistasis among nonallelic mutations.

General conditions and definitions

To begin, consider a population that is sufficiently large to ignore sampling error. The genome-wide mutation rate to mildly deleterious mutations is U_d. Mildly deleterious mutations are defined as those which, in the heterozygous state, reduce fitness by a few percent or less–most deleterious mutations fall into this category (Crow & Simmons, 1983). For simplicity, it is further assumed that each mutation reduces fitness by a constant decrement, *s*. Variation in *s* will not affect the major conclusions of this paper, although genetic drift will induce a stochastic form of mutational load whenever some mutations have minuscule effects (i.e. *s*-values less than $1/N_e$). Lastly, only forward mutations, from wild-type to mildly deleterious mutations, are explicitly modeled, although the consequences of rare reverse and compensating mutations are considered when relevant.

Genetic polarization

Sexual and asexual populations have a fundamental difference in their genetic structure that plays an important role in the accumulation of deleterious mutations (Charlesworth, 1994; Rice, 1996). Consider a population that has come to mutation-selection equilibrium with new deleterious mutations entering at rate U_d/genome/generation (Figure 1). When U_d is large (i.e., of the size presumed to occur in most eukaryotes), then most individuals will carry one or more deleterious mutations compared to the least-mutated class of individuals (Kimura & Maruyama, 1966).

When recombination (assumed throughout to include both segregation and intra-chromosomal recombination) is absent, the mutation process generates a unidirectional flow of new deleterious mutations from less to more mutated classes. This unidirectional flow is referred to as 'genetic polarization' (see for details, Rice, 1996). It dichotomizes an asexual population into: a) the progenitor class, i.e., the self-sustaining, least-mutated class, and b) all more heavily mutated classes. At equilibrium these more heavily mutated classes are not self-sustaining owing to the fact that each component mutational class relies for part of its reproduction on newly-mutated individuals cascading down from less mutated classes. Because they are not self-sustaining, these lineages are all marching toward eventual extinction, and hence they are collectively termed the 'living dead' (Rice, 1996). Because of genetic polarization, only the progenitor class gives rise to persistent genetic lineages.

Rare reverse and compensatory mutations can move deleterious mutations, via genetic hitchhiking, against the flow of genetic polarization. But this is a minor influence, analogous to water turbulence that occasionally transports a pebble a short distance upstream.

When recombination is present it moves new deleterious mutations bi-directionally to genetic backgrounds of higher and lower fitness. Because genes are no longer trapped in specific genetic backgrounds, all mutational classes can contribute to lineages that persist over evolutionary time.

Requisite load

Because of genetic polarization, only the fate of the self-sustaining progenitor class is relevant to an asexual population. But the fitness of the progenitor class depends on the level of competition from the living dead. So a measure of genetic load is developed that focuses on the progenitor class but incorporates the level of competition from the living dead.

This is done by integrating the necessary conditions for mutation-selection balance into the traditional definition of mutational load (*L*). For reasons stated previously, especially the complication of the unknown fitness of the unmutated class, the measure *L* has limited utility. Nonetheless we do know that the rate of removal of mutations from the gene pool is proportional to the standing variance in fitness (see below), so some requisite range in fitness $[\overline{W}_{best} - \overline{W}_{mean}]$ must be attained for mutation-selection balance to be achieved. We can therefore define the Requisite Standing Mutational Load (L_{req}) as the range in fitness, relative to the most fit extant genotype, needed to produce sufficient variance in fitness to halt the deterministic accumulation of mutations, i.e.,

$$L_{req} = [\overline{W}_{best(req)} - \overline{W}_{mean}],$$

where $\overline{W}_{best(req)}$ is the fitness of the best extant mutational class necessary to prevent deterministic mutation accumulation. This new measure of mutational load represents a genome-wide extension to an index previously proposed by Crow (1970; his measure III) in the context of a single haploid locus.

If we measure fitness in absolute terms, i.e., as the expected net reproductive rate (*R*) of different genotypes, then this requisite load equation becomes

$$L_{req} = R_{best(req)} - 1,$$

because the genetic equilibrium mean fitness ($\overline{R}$) must be unity at demographic equilibrium. A general solution for $R_{best(req)}$ is presented below, after a standard for comparison is established.

Dominant-sterile benchmark

To apply the concept of requisite load to the phenomenon of deterministic mutation accumulation, a benchmark is first established where mutational load is simple to calculate and its demographic interpretation is intuitively clear. Consider a population (sexual or asexual) that is semelparous and make the simplifying assumption that every mutation is dominant and causes sterility, i.e., individuals expressing one or more mutations are demographically and ecologically equivalent to unmutated individuals but produce no surviving offspring (offspring die immediately after the termination of parental investment).

The dominant-sterile mutations dichotomize the population into a fertile best-class (zero-mutations-class) and all remaining sterile classes carrying one or more mutations (this dichotomy is analogous to the progenitor class vs. the living dead). Assuming that new mutations are distributed as a Poisson variate among zygotes, then only a fraction e^{-Ud} (i.e., the size of the zero-class of a Poisson variate) of the zygotes will not receive a new sterile mutation, and the net reproductive rate of these unmutated individuals, required for the population to persist, must be $R_{best(req)} = e^{Ud}$, because only a fraction e^{-Ud} do not receive a new sterile mutation. Thus ($L_{req} = e^{Ud} - 1$) is the increment by which R_{best} must exceed unity to compensate for the load of deleterious mutations. This 'dominant-sterile' context will serve as a benchmark for mutational load under more realistic conditions.

Requisite load of an asexual population

Next we consider the mutational load of an asexual population experiencing recurrent mildly deleterious mutations. At any point in time only the progenitor class (best class) is generating persistent lineages, so we can focus exclusively on this subpopulation. Recall that a fraction $1 - e^{-Ud}$ of the offspring from the progenitor class receive one or more new mildly deleterious mutations and therefore do not contribute to this class next generation. Because these mutated zygotes are recruited to the living dead, they are effectively sterile because many will survive and act as competitors for resources. Thus the load of an asexual population is equivalent to that of the dominant sterile benchmark described above, i.e., $L_{req} = e^{Ud} - 1$. Whenever demographic, ecological, and/or physiological constraints cause R_{best} to be less than e^{Ud}, then the

progenitor class will decline in size each generation and deterministic mutation accumulation will ensue. Such mutation accumulation will be opposed by reverse and compensatory mutations, but if R_{best} is much less than e^{Ud}, then net mutation accumulation will ensue.

Accurate estimates of U_d for most metazoans with large genomes (e.g., vertebrates) are not yet available but extrapolations from studies of humans and *Drosophila* (Mukai, 1979; Kondroshov, 1988; Crow, 1993) suggest that $U_d > 5$ is feasible. In this case an asexual population would have to have $R_{best} > e^5 =$ 148, a value far beyond the physiological/ecological capabilities of most vertebrates. Thus once a species evolves to be sufficiently complex (i.e., with large genome and hence large U_d), then asexual reproduction is an evolutionary dead-end because it will lead to deterministic, open-ended mutation accumulation and eventual extinction.

Requisite load of a sexual population

In sexual populations, the best-class does not clonally reproduce itself but instead is produced via recombination and segregation of mildly deleterious mutations from the population as a whole. For example, if a member of the best-class mated at random, then very few of its offspring would be recruited to the best-class, because most offspring would carry an increased number of mutations. When recombination builds the best-class faster than its own net reproductive rate, then the requisite load of a sexual population is reduced.

The necessary conditions to prevent deterministic mutation accumulation in sexual vs. asexual populations can be compared by solving for the requisite net reproductive rate of the best-class ($R_{best(req)}$) in a sexual population. At equilibrium,

$$\begin{aligned}
\Delta\overline{n}_{sel} &= \overline{n}_{after\ selection} - \overline{n}_{prior\ to\ selection} \\
&= \left[\sum_{n=0}^{\infty} n[f(n)R_n/\overline{R}]\right] \quad - \overline{n} \qquad (1a) \\
&= [COV[n, R_n]/\overline{R}] \qquad (1b) \\
&= (B_{R,n})(\sigma_n^2)]/\overline{R} \qquad (1c)
\end{aligned}$$

where R_n and *f*(*n*) are the net reproductive rate and the fraction of individuals carrying *n* mildly deleterious mutations, respectively, $\Delta\overline{n}$ is the per generation change in the mean number of mutations per genome, $\overline{R}$ is the net reproductive rate of the population as a whole, σ_n^2 is the variance in *n*, *COV* denotes covariance, and $B_{R,n}$ is the regression (i.e. slope or selection gradient) of R_n on *n*. At equilibrium $\Delta\overline{n}_{sel} = -\Delta\overline{n}_{mut} = -U_d$, so,

$$\begin{aligned}
0 = \Delta\overline{n} = \Delta\overline{n}_{mut} + \Delta\overline{n}_{sel} = \\
U_d + [(\hat{B}_{R,n})(\hat{\sigma}_n^2)]/\hat{\overline{R}}
\end{aligned} \qquad (2)$$

where the superscript ∧ denotes the equilibrium value. If we next define w_n as the fitness of genotypes relative to the most fit extant genotype, then $w_n = R_n/R_{best}$. Assuming demographic equilibrium, $\hat{\overline{R}}$ = unity and rearrangement of (2) yields,

$$R_{best(req)} = 1/\left[(-\hat{B}_{w,n})(\hat{\sigma}_n^2/U_d)\right], \qquad (3)$$

where $\hat{B}_{w,n}$ is the equilibrium regression of relative fitness (w_n) on *n*.

Equation (3) can be made more intuitive by expressing its right side in terms of the selection gradient and opportunity for selection of the dominant-sterile benchmark. Note that $\hat{B}_{w,n}$ is maximized in the most extreme case of the dominant-sterile benchmark. In this case, $\hat{B}_{w,n} = -e^{-Ud}$ (Appendix 1). This motivates the index = $\hat{S}^*$ = standardized selection gradient = (observed selection gradient) / (its maximal possible value) = $\hat{B}_{w,n}/(-e^{-Ud})$.

To simplify further, note that the last factor in equation (3), $(\hat{\sigma}_n^2/U_d)$, is the ratio of the standing variance ($\hat{\sigma}_n^2$) in the number of mutations per individual (*n*) divided by its minimal value, i.e., divided by the mutational variance in *n* (U_d), which is the standing variance for the case of the dominant-sterile benchmark. This ratio can be used as a standardized measure of the relative opportunity for selection to act and is denoted here by $\hat{O}^*$. Substituting $\hat{S}^*$ and $\hat{O}^*$ into (9),

$$R_{best(req)} = e^{Ud}/(\hat{S}^*\hat{O}^*). \qquad (4)$$

Thus the requisite net reproductive rate of the best-class is a simple function of the standardized selection gradient and opportunity for selection.

When there is neither epistasis nor sampling error, linkage disequilibrium will be absent, recombination will have no effect, and sexual and asexual populations will have identical genotypic distributions (Kimura & Maruyama, 1966; Haigh, 1978). Therefore, at equilibrium, $R_{best(req),sexual} = R_{best(req),asex} = e^{Ud}$.

Because genetic polarization necessarily causes the load of an asexual population to be $e^{Ud} - 1$, it follows that $\hat{S}^*\hat{O}^* \equiv 1.0$ in the absence of recombination. But as shown below, in sexually recombining populations,

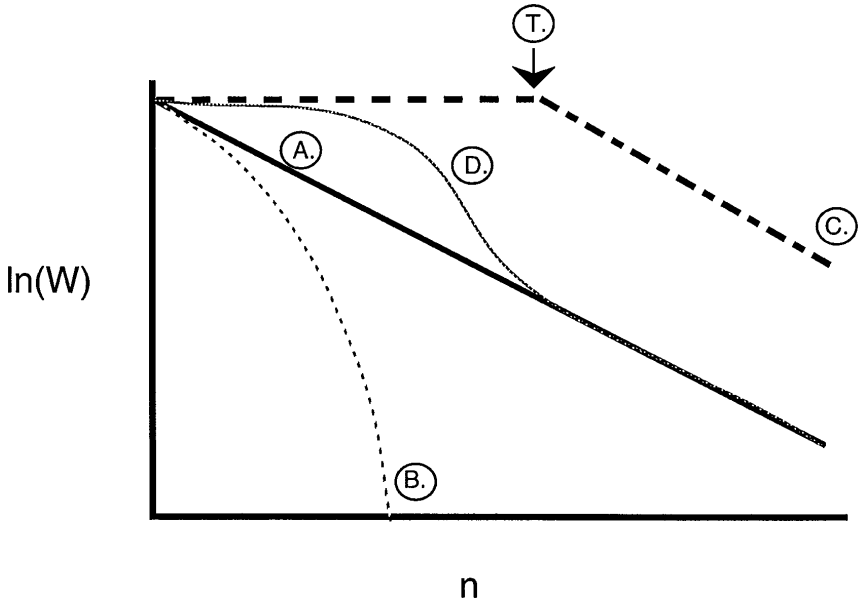

Figure 2. A plot of the log of fitness vs. the number of mutations in a genome illustrates major forms of epistasis: *A:* no epistasis; *B:* reinforcing epistasis; *C:* buffering epistasis; *T:* threshold for buffering epistasis; *D:* incomplete buffering epistasis.

epistasis and nonrandom mating can cause the standardized selection gradient ($\hat{S}^*$) and the opportunity for selection ($\hat{O}^*$) to change in a non-compensating manner and thereby cause the requisite load to deviate from $e^{Ud} - 1$. When the product $\hat{S}^*\ \hat{O}^*$ increases beyond unity, then recombination builds the best-class faster than its own net reproductive rate. Thus recombination can increase the efficiency of selection and thereby reduce the mutational load. This motivates the index $Eff^* = \hat{S}^*\hat{O}^*$ as an index of the relative efficiency of selection.

Impact of epistasis and nonrandom mating on mutational load

General

When epistasis is absent, selection acts independently on each mutation and fitness declines linearly, on a log scale, with the number of mutations in a genome (Figure 2A). Epistasis among mutations represents synergism (i.e., the total effect is greater than the sum of the component effects when acting independently) and it can be positive or negative. Diminishing-returns epistasis is the major negative form. In this case, the deleterious impact of a mutation decreases with each additional mutation in its genetic background. It is well established that diminishing-returns epistasis increases mutational load when recombination is present (Crow, 1970) and it will not be discussed further here.

Two types of positive epistasis have traditionally been discussed concerning a reduced mutational load of a sexual population, i.e., that associated with reinforcing and truncation selection. Reinforcing epistasis occurs when the deleterious effect of a mutation increases with each additional mutation in its genetic background (Figure 2B). Truncation selection, in its simplest form, occurs when mutations do not reduce fitness until they accumulate beyond a threshold (*T*), at which point fitness plummets to zero. Truncation selection combines two forms of epistasis: a) buffering epistasis, whereby the harm of individual deleterious mutations is ameliorated except for those in excess of a threshold value (*T*), and b) reinforcing epistasis once the threshold is reached. Because truncation selection includes reinforcing epistasis, the focus here will only be on the independent effects of reinforcing (Figure 2B) vs. buffering epistasis (Figure 2C).

This will be done by examining the impact of these two forms of epistasis on the efficiency of selection (Eff^*) and its components ($\hat{S}^*$ and $\hat{O}^*$). Because the direct effect of recombination is to reduce linkage disequilibrium, we also need a measure of the impact of recombination on linkage disequilibrium. This is

done by expanding the expression for the opportunity for selection as

$$\hat{O}^* = \hat{\sigma}_n^2/U_d = [\hat{\overline{n}}/U_d](\hat{D}^*) = \hat{A}^*\hat{D}^*,$$

where $\hat{D}^* = \hat{\sigma}_n^2/\hat{\overline{n}}$ and $\hat{A}^* = \hat{\overline{n}}/U_d$.

The parameter $\hat{A}^*$ is the standardized accumulation of mutations, i.e., the actual accumulation ($\hat{\overline{n}}$) divided by this value in the case of the dominant sterile benchmark (U_d). In the absence of linkage disequilibrium, an increase in the accumulation of mutations ($\hat{A}^*$) produces a corresponding increase in the opportunity for selection ($\hat{O}^*$).

The parameter $\hat{D}^*$ measures linkage disequilibrium in units of $\hat{\sigma}_n^2$. If there is a random association between mutations (i.e., no linkage disequilibrium), then *n* follows a Poisson distribution and $\hat{D}^*$ = 1 (i.e., mean = variance for a Poisson variate). $\hat{D}^*$ measures the degree to which a nonrandom association between mutations increases or decreases the $\hat{\sigma}_n^2$. Requisite load can now be expressed as a simple function of the equilibrium level of linkage disequilibrium,

$$R_{best(req)} = e^{Ud}/(\hat{S}^*\hat{A}^*\hat{D}^*). \qquad (5)$$

Lastly, the measure $ln(D^*)$ will also be used as a measure of linkage disequilibrium since, on a log scale, negative values of D^* reduce, and positive values increase, the opportunity for selection.

Simulation model

A deterministic simulation model, originally described by Kimura and Maruyama (1966), was used to explore the impact of buffering and reinforcing epistasis on the efficiency of selection in sexual and asexual populations. The model simulates free recombination between all loci. Two parameters were constant across all simulations: $U_d = 1.0$ and $s = 0.05$ (recall that *s* is the decrement to fitness when a single mutation is expressed).

The absence of epistasis was modeled by the fitness function, $w_n = (1-s)^n$. Buffering epistasis was modeled by the fitness function; $w_n = 1$ if $n \leq T$, and $w_n = (1-s)^{n-T}$, if $n > T$, where the threshold for buffering epistasis (*T*; Figure 2C) is; $T = E(n \mid \text{no epistasis}) = U_d/s = 20$ (Haig, 1978). Reinforcing epistasis was modeled by the fitness function; $w_n = (1-s)^{f(n)}$, where $f(n) = n^{1.5}$. These parameter values are arbitrary, but the qualitative relationships are unchanged when other parameter values are used.

To display the effects of epistasis in asexual and recombining populations, graphs were constructed showing the change in the equilibrium values of $\hat{S}^*$, $\hat{O}^*$, $\hat{D}^*$, $\hat{A}^*$, and Eff^*, both when recombination was present and absent. Because these parameters combine multiplicatively, all changes are expressed as proportions on a log scale. These proportional changes will be expressed as '*ln*(Delta)' values. For example, if epistasis increased $\hat{O}^*$ by 10% (i.e., from 1.0 to 1.1) then *ln*(Delta) = *ln*(1.1/1.0), and if it reduced $\hat{O}^*$ by 10% then *ln*(Delta) = *ln*(0.9/1.0).

The base for the graphs in Figure 2 is the case where epistasis is absent (not shown in Figure 2). Here the *ln*(Delta) values are zero for all of the parameters. The graphs display the impact of epistasis (or nonrandom mating) on each parameter when recombination is present and when it is absent. Positive (negative) histograms indicate the proportional increase (decrease) in a parameter. Because of the multiplicative interaction between the parameters, the histograms for* $\hat{S}^*$ and $\hat{O}^*$ sum to that of Eff^*, and those of $\hat{A}^*$ and $\hat{D}^*$ sum to $\hat{O}^*$.

Simulation results

Figure 3A illustrates the effects of buffering epistasis. When recombination is present, it keeps the level of linkage disequilibrium small (ln($\hat{D}^*$) $\approx$ 0). The direct effect of buffering epistasis is to increase the accumulation of mutations ($\hat{A}^*$), and this in turn increases the opportunity for selection ($\hat{O}^*$). Epistasis causes more mutations to accumulate, because selection acts only on the number of mutations in excess of the threshold number (*T*).

Interestingly, buffering epistasis also has a large indirect effect: it makes the equilibrium selection gradient steeper. The value of $\hat{S}^*$ increases because the σ_n^2 per expressed mutation (i.e., number mutations in excess of the buffering threshold, *T*) is increased. This in turn reduces the equilibrium mean number of expressed mutations (i.e., $\hat{\overline{n}}_{expressed} = \hat{\overline{n}} - T$), compared to the case of no epistasis (i.e., $\hat{A}^*$ is increased but not $\hat{\overline{n}}_{expressed}$). All else being equal, the impact of an arbitrary individual mutation on fitness is greater when fewer mutations are present (and expressed) in the genome. For example, consider *s* = 0.05 when buffering epistasis is not operating. With one expressed mutation, w_1 = (1–.05) and the single mutation reduces fitness by 5%. But if 20 mutations are present, w_{20} = (1–

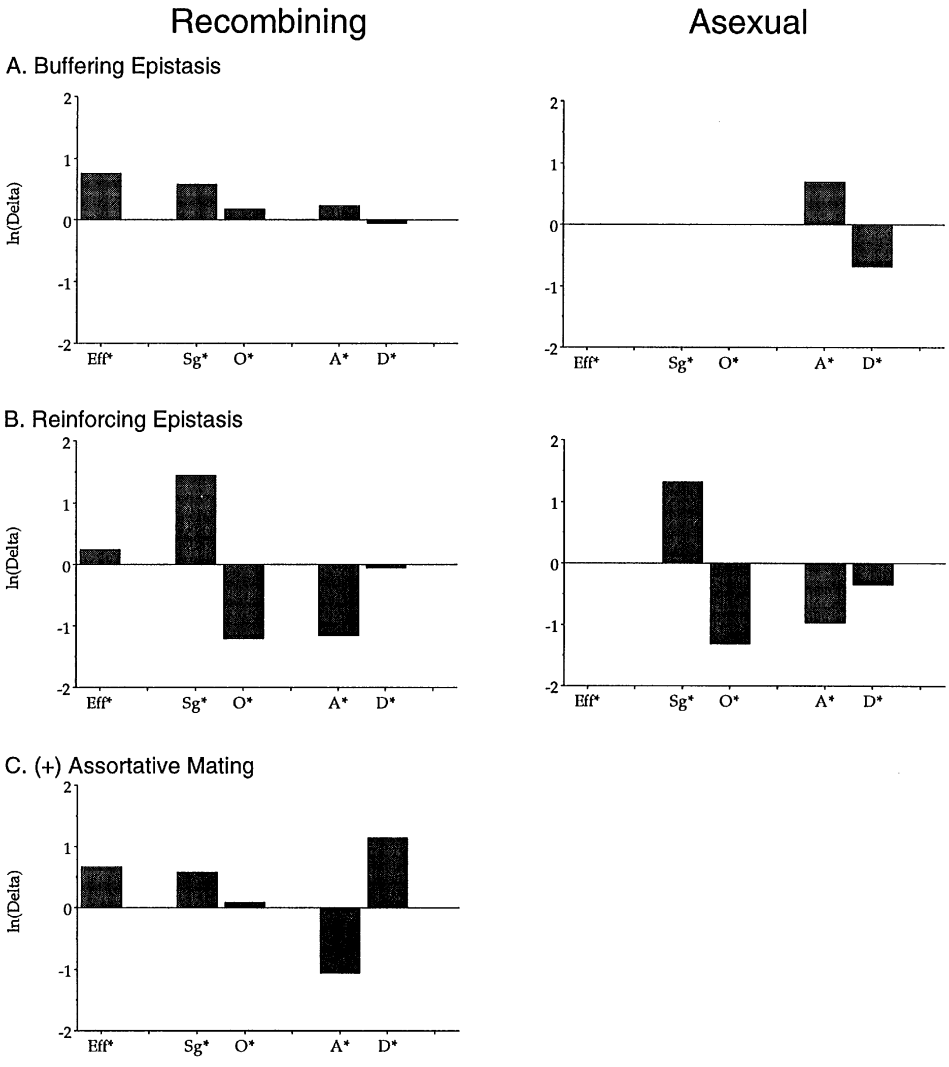

Figure 3. The effects of epistasis, positive assortative mating, and sex vs. asex on the standardized opportunity for selection ($\hat{O}^*$), the standardized selection gradient ($\hat{S}^*$) linkage disequilibrium ($\hat{D}^*$), and the efficiency of selection (*Eff**). When there is no epistasis, the values $\hat{S}^*$, $\hat{O}^*$, $\hat{D}^*$, $\hat{A}^*$, and *Eff** necessarily equal unity, and this transforms to zero on the natural log scale of the graphs ($\hat{S}^*$ is Sg^* on figure).

.05)^{20} = .3585 and each mutation, on average, reduces fitness by only 3.2%.

When recombination is absent, buffering epistasis builds negative (variance reducing) linkage disequilibrium ($ln[\hat{D}^*] < 0$). This prevents the increase in $\hat{A}^*$ from yielding a net increase in $\hat{O}^*$. The build-up of this linkage disequilibrium drives $\hat{S}^*$, $\hat{O}^*$, and Eff^* all to unity and thus $L_{req} = e^{Ud} - 1$, i.e., the same as the dominant-sterile benchmark. Thus when recombination is absent, epistasis builds 'compensatory disequilibrium,' and $\hat{D}^*_{asex} = 1/(\hat{S}^*\hat{A}^*)$ and $ln(\hat{D}^*_{asex}) = -[ln(\hat{S}^*) + ln(\hat{A}^*)]$.

The buffering of deleterious mutations can be incomplete but still provide a large advantage. To simulate incomplete buffering epistasis (Figure 2D) I fit the fitness model: $w_n = [(1-s)^n]^{f(n,T)}$, where $f(n,T) = (n/T)^2$ if $n < T$ and $w_n = (1-s)^n$ if $n \geq T$. In this case most (83%) of the advantage of complete buffering epistasis was realized. L_{req} is 1.72 for asex, and 0.53 for incomplete and 0.29 for complete buffering epistasis.

Figure 3B illustrates the effects of reinforcing epistasis. When recombination is present, linkage disequilibrium $[ln(\hat{D}^*)]$ is kept near zero. The selection

gradient ($\hat{S}^*$) is increased directly due to reinforcing epistasis increasing the strength of selection. $\hat{S}^*$ is also increased indirectly due to a lowered equilibrium value of $\hat{\overline{n}}$ (i.e., reduced $\hat{A}^*$, see above). The opportunity for selection ($\hat{O}^*$) is diminished because the stronger selection (due to reinforcing epistasis) reduces $\hat{A}^*$. But because the selection gradient ($\hat{S}^*$) is increased more than the opportunity for selection ($\hat{O}^*$) declines, the net effect is an increase in the efficiency of selection (*Eff**) and hence a diminished mutational load ($L_{req} < e^{Ud} - 1$).

When recombination is absent, reinforcing epistasis builds compensatory disequilibrium. This variance-reducing, negative disequilibrium reduces $\hat{O}^*$ to the point that the gain in the selection gradient is exactly off-set by the decline in the opportunity for selection.

In addition to epistasis, nonrandom mating can also reduce the requisite mutational load. Figure 3C illustrates the impact of positive assortative mating (individuals mate only with others carrying the same number of mildly deleterious mutations). This form of nonrandom mating causes the σ_n^2 to rise far more rapidly with increasing $\overline{n}$, and thereby increases $\hat{O}^*$ despite the reduction in the accumulation of mutations ($\hat{A}^*$). The reduction in $\hat{A}^*$ increases the selection gradient ($\hat{S}^*$), as described above for the case of reinforcing epistasis. Unlike the two forms of epistasis examined above, however, positive assortative mating builds strong beneficial (i.e., variance enhancing) linkage disequilibrium $[ln(\hat{D}^*) >> 0]$ and can thereby provide a substantial reduction in the requisite mutational load. Even weak positive assortative mating can produce a large reduction in load (data not shown). Obviously, negative assortative mating would increase the requisite load of a recombining population.

Overall, epistasis and nonrandom mating can cause recombination to build the best-class faster than its own net reproductive rate. Clearly when $R_{best(realized)} << R_{best(req)}$, then deterministic mutation accumulation will lead to extinction. But when the increment ($R_{best(req)} - R_{best(realized)}$) is smaller, a mutational Red-Queen may ensue with mutation accumulation being accommodated by perpetual compensating adaptation. Recombination, by reducing $R_{best(req)}$ via epistasis and nonrandom mating, extends the permissible range of phenotypic complexity (large U_d) that can potentially evolve.

Pathway-epistasis and load reduction

Is epistasis strong enough in natural populations to provide a non-trivial reduction in the requisite mutational load of sexually recombining populations? There is some evidence for weak reinforcing epistasis from studies that examined the fitness of genotypes containing increasing numbers of random mutations (i.e., heterozygous recessive lethals) when stacked into the same genome (Kitagawa, 1967) or when increasing numbers of random mutations are accumulated on sheltered chromosomes (reviewed in Mukai, 1979). But studies such as these are severely biased against the possibility of detecting strong epistasis, because they are concerned with the combined effects of random mutations that are unlikely to interact directly with one another.

A more powerful approach would be to look at closely interacting loci, such as those involved, for example, in enzymatic, signal transduction, and developmental pathways. Such groups of functionally interacting loci may be of major importance since only a small minority of all possible combinations of genome-wide mutations need to interact epistatically to provide a large advantage to sexual recombination, i.e., $L_{req} << e^{Ud} - 1$.

To see why this is the case, suppose that a genome could be decomposed into *K* independent pathways of strongly interacting loci. If buffering and/or reinforcing epistasis were expressed among mutations within, but not between, pathways then

$$R_{best(req)} = e^{Ud} / \prod_{i=1}^{K} \hat{S}_i^* \hat{O}_i^*, \qquad (6)$$

where the subscript *i* denotes a specific pathway and with the constraint that $\hat{S}_i^* \hat{O}_i^* \leq e^{(Ud)i}$ (Appendix 2). For example suppose U_d = 2.0 and there are 100 independent pathways with: 1) equal portions of the genome associated with each pathway, 2) no epistasis between mutations from different pathways, and 3) buffering epistasis within pathways which increased $\hat{S}_i^* \hat{O}_i^*$ from unity to 1.0186 (Appendix 3). The requisite net reproductive rate of the least mutated class ($R_{best(req)}$) declines from 7.39, assuming no pathway-epistasis or asexual reproduction, to only 1.16.

Evidence for pathway epistasis

Evidence for strong epistasis within pathways is still fragmentary. Control theory predicts that buffering epistasis may be an incidental byproduct of the inherent properties of flux through enzymatic pathways (Kacser & Burns 1979; Szathmary, 1993).

This theory, however, is based on extremely simplified conditions. Most pathways operating in nature will be forced to operate under a wide range of cytosolic conditions, including factors such as desiccation, salt imbalance, broadly varying temperatures across both time and position within the organism, trace element deficiency/surplus, heat-shock induced deficits in availability of constituent enzymes, and many other factors too numerous to list here. This variation in cytosolic conditions in the past should have selected for resilient pathways. A recent theoretical analysis of a signal transduction pathway in bacteria supports this conclusion (Barkai & Leibler, 1997). If pathways have evolved that are robust to such environmental insults, then those same pathways may be preadapted to tolerate the minor genetic insults represented by mildly deleterious mutations. Thus past selection for pathways to operate under a wide range of environmental conditions may have fortuitously built buffering epistasis among strongly interacting mutations.

One clear example of buffering epistasis within a developmental pathway comes from the early work on genetic assimilation and canalization (e.g., Waddington, 1953). For example, the crossveinless phenotype is extremely rare in laboratory populations of *D. melanogaster*, but this condition is seen at low frequency when the flies are heat shocked during development. By selecting, over many generations, only those flies that expressed the crossveinless phenotype under heat shock, Waddington generated lines that expressed the trait without heat shock (genetic assimilation). Thus the genetic variation for the crossveinless phenotype was initially present at low frequency but hidden by canalization (buffering epistasis) when rare. Once the genetic variation accumulated sufficiently (beyond threshold levels), it was expressed. This and other studies on canalization provide preliminary evidence for epistasis in the context of developmental pathways.

Conclusions and future research

Load, sex, and epistasis

A major goal of this paper was to develop a heuristic model that illustrates *how* sex (recombination and mating) can reduce mutational load. Inspection of the expression for requisite load,

$$L_{req} = R_{best(req)} - 1 = \{e^{Ud}/[\hat{S}^*\hat{O}^*]\} - 1 = \{e^{Ud}/[\hat{S}^*\hat{A}^*\hat{D}^*]\} - 1$$

makes it clear that the critical feature of sex is to break the antagonism (i.e., compensating changes) between the equilibrium selection gradient ($\hat{S}^*$) and opportunity for selection ($\hat{O}^*$). When reinforcing and/or buffering epistasis is present, this antagonism is mediated by the build-up of compensatory linkage disequilibrium.

Beneficial epistasis is a two-edged sword; it helps by increasing the selection gradient and/or the opportunity for selection, but it hurts by inducing the build-up of compensatory linkage disequilibrium. Asexual populations experience both the advantage and disadvantage of epistasis, and consequently experience no net gain. Recombination, however, rescues sexual populations from the disadvantage of compensatory disequilibrium.

The parameters $\hat{S}^*$ and $\hat{O}^*$ have simple intuitive interpretations. In a general sense, mutational load is reduced when selection culls more mutations per selective death (or sterility). Thus the more efficient selection is in removing mutation-rich genotypes, in excess of mutation-poor genotypes, the lower the mutational load.

The selection gradient measures the degree to which selection culls more heavily from mutation-rich compared to mutation-poor genotypes (i.e., from the right as opposed to the left side of the fitness distribution of Figure 1). All else being equal, the greater the selection gradient the greater the number of mutations removed per selective death (or sterility), and hence the greater the efficiency of selection (*Eff**). But a steep selection gradient has a cost in the following generation, because there will be fewer mutation-rich genotypes available (negative linkage disequilibrium; $ln(D^*) < 0$) and this reduces the opportunity for selection.

The opportunity for selection measures the degree to which a population regenerates mutation-rich genotypes. With asexual reproduction, only the weak force of mutation regenerates the genotypes removed by

selection. When epistatic selection removes mutation-rich genotypes faster than mutation can regenerate them, compensatory (i.e., negative) disequilibrium accrues and $\hat{O}^*$ declines. With sex, recombination rapidly regenerates the mutation-rich genotypes removed by epistatic selection (moves $ln[\hat{D}^*]$ toward 0) and the advantage of epistasis is realized.

In summary, genetic polarization isolates the progenitor class from the remainder of the population. As a result, any detrimental mutation (no matter how minor) is effectively a dominant sterile, and this constrains the requisite load of an asexual population to be $L_{req} = R_{best(req)} - 1 = e^{Ud} - 1$. Since the fitness of the progenitor class depends upon the level of competition from the living dead (i.e., $R_{best(req)} = 1/E(w_n)$; Appendix 2), this constraint can be expressed as the build-up of population-wide compensatory linkage disequilibrium when epistasis is present, i.e., $ln(\hat{D}^*_{asex}) = -[ln(\hat{S}^*) + ln(\hat{A}^*)]$. But when sexual recombination is present, the build-up of compensatory disequilibrium is suppressed, $\ln(\hat{D}^*_{sex}) \approx 0$. In this case: (a) Reinforcing epistasis reduces mutational load primarily by steepening the selection gradient. This causes selection to cull more heavily from the mutation-rich side of the fitness distribution. (b) Buffering epistasis permits more mutations to accumulate. This makes all individuals more mutation-rich without a concomitant reduction in the selection gradient. (c) Positive assortative mating increases the frequency of extreme genotypes (i.e., the relative proportion of both mutation-rich and mutation-poor individuals). As a consequence a greater fraction of mutation rich genotypes are accessible to selection. In all three cases, recombination causes selection to cull more mutations per selective death (or sterility) and mutational load is reduced.

New research

The second goal of this paper was to redirect the focus of epistasis research, in the context of the adaptive significance of recombination, from genome-wide to pathway epistasis. Inspection of the expression for L_{req} makes it clear that a powerful way of reducing mutational load is to increase the load of mutations ($\hat{A}^*$) without concomitantly decreasing the selection gradient ($\hat{S}^*$). Genome-wide truncation selection is an obvious way of doing this, but buffering epistasis within pathways has equal potential to reduce mutational load. While there is theory to support the operation of such buffering epistasis within pathways (Barkai & Leibler, 1997), experimental support is critically lacking.

Because so many pathways have been well-characterized by molecular biologists, there is a unique opportunity for biochemists and physiologists to contribute to our understanding of the adaptive significance of sexual recombination. Studies in which minor effect mutations are stacked within pathways could be used as a powerful test for buffering and reinforcing epistasis. A preliminary study concerning a developmental pathway is currently underway in my laboratory. Ecological studies of the extent of positive assortative mating for fitness also seem feasible, but to my knowledge, are absent.

References

Barkai, N. & S. Leibler, 1997. Robustness in simple biochemical networks. Nature 387: 913–916.

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

Crow, J.F., 1970. Genetic loads and the cost of natural selection, pp. 128–177 in Mathematical Topics in Population Genetics, Volume 1, edited by K. Kojima. Springer-Verlag, New York.

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

Crow, J.F. & M.J. Simmons, 1983. The mutational load of *Drosophila*, pp. 1–35 in The Genetics and Biology of *Drosophila*, edited by M. Ashburner, H.L. Carson, & J.N. Thompson. Academic Press, New York.

Haigh, J., 1978. The accumulation of deleterious genes in a population – Muller's ratchet. Theoret. Pop. Biol. 14: 251–267.

Mukai, T., 1979., Polygenic mutations, pp. 177–196 in Quantitative Genetic Variation, edited by N.J. Thompson, Jr. and J.M. Thoday. Academic Press, New York.

Kacser, H. & J.A. Burns, 1979. Molecular democracy: who shares the controls. Biochemical Society Transactions 7: 1149–1160.

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

Kitagawa, O., 1967. Interactions in fitness between lethal genes in heterozygous condition in *Drosophila melanogaster*. Genetics 57: 809–820.

Kondroshov, A.S., 1988. Deleterious mutations and the evolution of sexual recombination. Nature 336: 435–440.

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

Szathmary, E., 1993. Do deleterious mutations act synergistically? Metabolic control theory provides a partial answer. Genetics 133: 127–132.

Waddington, C.H., 1953. Genetic assimilation of an acquired character. Evolution 7: 118–126.

Appendix 1

The regression ($B_{w,n}$) of relative fitness (w_n) on the number of mutations per genome (n) is equal to the $COV(w_n, n)/\hat{\sigma}_n^2$. In the dominant-sterile case, w_n is dichotomous, equaling 1 when $n = 0$ and 0 otherwise. Because the distribution of new mutations is assumed to be Poisson (U_d) and the because the probability density at zero of such a Poisson variate is e^{-Ud}, the expectation of w_n is

$$E(w_n) = \sum_{n=0}^{\infty}(w_n)f(n) = 1(e^{-Ud}) + 0 + 0 + ... = e^{-Ud}.$$

Because the expectation and variance of Poisson variate are equal, the $E(n) = \sigma_n^2 = U_d$. With the probability density of $n, E(w_n)$, and $E(n)$ defined, the covariance between n and w_n can be expressed as

$$COV(w_n, n) = \left[\sum_{n=0}^{\infty} n(w_n)f(n)\right] - E(n)E(w_n) = 0 - E(n)E(w_n)$$

because all terms in the summation are zero. Next, the regression of w_n on $n(B_{w,n})$ can be expressed as

$$\begin{aligned} B_{w,n} &= COV(w_n, n)/\sigma_n^2 = -(U_d)(e^{-Ud})/U_d \\ &= -e^{-Ud}. \end{aligned}$$

Appendix 2

To begin, consider the definition of w_n at mutation-selection balance,

$$w_n = R_n/R_{best(req).}$$

Taking expectations of both sides,

$$\begin{aligned} E(w_n) &= E(R_n/R_{best(req).}) \\ &= (1/R_{best(req).})E(R_{n.}) \\ &= 1/R_{best(req)}, \end{aligned}$$

because the $E(R_n) = 1$ at demographic equilibrium. By rearrangement, $R_{best(req)} = 1/E(w_n)$. If there are K independent pathways and no epistasis between mutations residing in different pathways, then

$$\begin{aligned} E(w_n) &= \prod_{i=1}^{K} E(w_n)_i \\ &= \sum \prod_{i=1}^{K} (e^{-(Ud)\,i}\hat{S}_i^*\hat{O}_i^*) \\ &= e^{-Ud}\prod_{i=1}^{K}\hat{S}_i^*\hat{O}_i^*. \end{aligned}$$

Finally, since $R_{best(req)} = 1/E(w_n)$,

$$R_{best(req)} = e^{Ud}/\prod_{i=1}^{K}\hat{S}_i^*\hat{O}_i^*$$

Appendix 3

Assuming $U_d = 2$ and that there are 100 independent pathways, then in a eukaryote with more than 10,000 loci per genome there would be more than 100 interacting loci per pathway. Next suppose that buffering epistasis within pathways suppresses mutation expression until at least 30 have accumulated (i.e, $w_n = 1.0$ when $n<30$), at which point each additional mutation reduces fitness by 5%, $(w_n = [1 - 0.05]^{n-25})$. Using the deterministic simulation model described in the text with $U_d = 0.02$, then $\hat{S}_i^*\hat{O}_i^*$ was found to equal 1.0186.