Rate of Information Acquisition by a Species subjected to Natural Selection

David J.C. MacKay
Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE. United Kingdom.
mackay@mrao.cam.ac.uk

May 19, 1999 — Draft 5.3

Abstract

At what rate, in bits per generation, can the blind watchmaker cram information into a species by natural selection? And what is the maximum mutation rate that a species can withstand? We study a simple model of a reproducing population of N individuals with a genome of size G bits: fitness is a strictly additive trait subjected to directional selection; variation is produced by mutation or by recombination and truncation selection selects the N fittest children at each generation to be the parents of the next. We find striking differences between populations that have recombination and populations that do not. If variation is produced by mutation alone, then the entire population gains up to roughly 1 bit per generation. If variation is created by recombination, the population can gain $O(\sqrt{G})$ bits per generation. Furthermore, recombination raises the maximum mutation rate that can be tolerated by a factor of order $\sqrt{G}$. This model explains the prevalence of sex in evolution and shows why sex persists in species with large genomes, even when they have reached evolutionary stasis.

1 The Model

It has been suggested that the deleterious mutation rate in hominids is so high that sex is essential for such species to persist (Eyre-Walker and Keightley, 1999). In order to quantify the benefit of sex, we consider a simple model of an evolving population of N individuals having a genome of size G. This 'threshold selection' model has previously been used in discussions of the rate of evolution by Maynard Smith (1968). We choose a crude model because it readily yields simple but striking results.

The genotype of each individual is a vector $\mathbf{x}$ of G bits, each having a good state $x_g = 1$ and a bad state $x_g = 0$. The fitness $F(\mathbf{x})$ of an individual is simply the sum of her bits:

$$F(\mathbf{x}) = \sum_{g=1}^{G} x_g. \tag{1}$$

The bits in the genome could be considered to correspond either to genes that have good alleles ($x_g = 1$) and bad alleles ($x_g = 0$), or to the nucleotides of a genome. We will concentrate on the latter interpretation. The essential property of fitness that we are assuming is that it is locally a roughly linear function of the genome, that is, that there are many possible changes one could

make to the genome, each of which has a small effect on fitness, and that these effects combine approximately linearly.

We consider evolution by natural selection under two models of variation.

Variation by mutation: The model assumes discrete generations. At each generation, t, every individual produces two children. The children's genotypes differ from the parent's by random mutations. Natural selection selects the fittest N progeny in the child population to reproduce, and a new generation starts.

[The selection of the fittest N individuals at each generation is known as truncation selection.]

The simplest model of mutations is that the child's bits $\{x_g\}$ are independent. Each bit has a small probability of being flipped, which, thinking of the bits as corresponding roughly to nucleotides, is taken to be a constant m, independent of x_g. [If alternatively we thought of the bits as corresponding to genes, then we would model the probability of the discovery of a good gene, $P(x_g=0 \rightarrow x_g=1)$, as being a smaller number than the probability of a deleterious mutation in a good gene, $P(x_g=1 \rightarrow x_g=0)$.]

Variation by recombination (or crossover, or sex): Our organisms are haploid, not diploid. They enjoy sex by recombination. The N individuals in the population are married into $M = N/2$ couples, at random, and each couple has C children — with $C = 4$ children being our standard assumption, so as to have the population double and halve every generation, as before. The C children's genotypes are independent given the parents'. Each child obtains its genotype $\mathbf{z}$ by random crossover of its parents' genotypes, $\mathbf{x}$ and $\mathbf{y}$. The simplest model of recombination has no linkage, so that:

$$z_g = \begin{cases} x_g & \text{with probability } 1/2 \\ y_g & \text{with probability } 1/2 \end{cases} . \tag{2}$$

Once the MC progeny have been born, the parents pass away, the fittest N progeny are selected by natural selection, and a new generation starts.

We now study these two models of variation in detail.

1.1 Relationship between fitness increase and information acquisition

If the bits are set at random, the fitness is roughly $F = G/2$. If evolution leads to a population in which all individuals have the maximum fitness $F = G$, then G bits of information have been acquired by the species, namely for each bit x_g, the species has figured out which of the two states is the better.

If the species is in some intermediate state where a fraction f_g of the population has $x_g = 1$, we define the information acquired to be

$$I = \sum_g \log_2 \frac{f_g}{1/2} \text{bits}, \tag{3}$$

because $\log_2(1/f)$ is the information required to find a black ball in an urn containing black and white balls in the ratio $f : 1-f$. If all the fractions f_g are equal to F/G, then

$$I = G \log_2 \frac{2F}{G}, \tag{4}$$

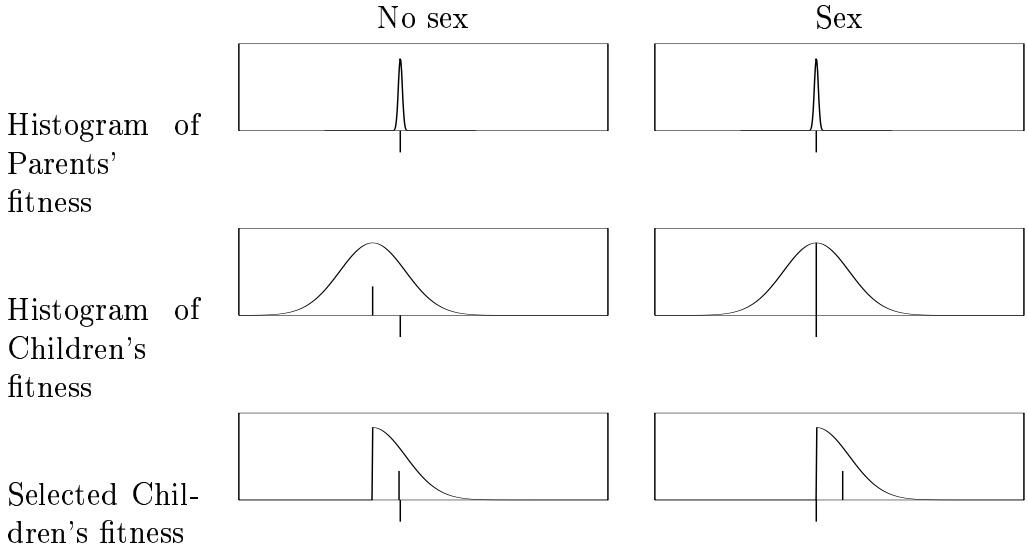

Figure 1: Explanation, in a nutshell, for the advantage of sex over parthenogenesis. If mutations are used to create variation among children, then it is unavoidable that the average fitness of the children is lower than the parents' fitness; the greater the variation, the greater the deficit. Selection bumps up the mean fitness again. In contrast, recombination produces variation without a decrease in average fitness. The typical amount of variation scales as $\sqrt{G}$, where G is the genome size, so after selection, the average fitness rises by $O(\sqrt{G})$.

which is well approximated by

$$\tilde{I} \equiv 2(F - G/2). \tag{5}$$

The rate of information acquisition is thus roughly two times the rate of increase of fitness in the population. We define the normalized fitness $f(\mathbf{x}) \equiv F(\mathbf{x})/G$.

2 Rate of increase of fitness

2.1 Theory of mutations

We assume that the genotype of an individual with normalized fitness $f = F/G$ is subjected to mutations that flip bits with probability m. We first show that if the average normalized fitness f of the population is greater than $1/2$, then the optimal mutation rate is small, and the rate of acquisition of information is of order one bit per generation.

Since it is easy to achieve a normalized fitness of $f = 1/2$ by simple mutation, we'll assume $f > 1/2$ and work in terms of the excess normalized fitness $\delta f \equiv f - 1/2$. If an individual with excess normalized fitness δf has a child and the mutation rate m is small, the probability distribution of the excess normalized fitness of the child has mean

$$\overline{\delta f}_{\text{child}} = (1 - 2m)\delta f \tag{6}$$

and variance

$$\frac{m(1-m)}{G} \simeq \frac{m}{G}. \tag{7}$$

If the population of parents has mean $\delta f(t)$ and variance $\sigma^2(t) \equiv \beta\frac{m}{G}$, then the child population, before selection, will have mean $(1-2m)\delta f(t)$ and variance $(1+\beta)\frac{m}{G}$. Natural selection chooses the upper half of this distribution, so the mean fitness and variance of fitness at the next generation are given by

$$\delta f(t+1) = (1-2m)\delta f(t) + \alpha\sqrt{(1+\beta)}\sqrt{\frac{m}{G}}, \tag{8}$$

$$\sigma^2(t+1) = \gamma(1+\beta)\frac{m}{G}, \tag{9}$$

where α is the mean deviation from the mean, measured in standard deviations, and γ is the factor by which the child distribution's variance is reduced by selection. The numbers α and γ are of order 1. For the case of a Gaussian distribution, $\alpha = \sqrt{2/\pi} \simeq 0.8$ and $\gamma = (1-2/\pi) \simeq 0.36$. If we assume that the variance is in dynamic equilibrium, *i.e.*, $\sigma^2(t+1) \simeq \sigma^2(t)$, then

$$\gamma(1+\beta) = \beta, \text{ so } (1+\beta) = \frac{1}{1-\gamma}, \tag{10}$$

and the factor $\alpha\sqrt{(1+\beta)}$ in equation (8) is equal to 1, if we take the results for the Gaussian distribution, an approximation that becomes poorest when the discreteness of fitness becomes important, *i.e.*, for small m. The rate of increase of normalized fitness is thus:

$$\frac{df}{dt} \simeq -2m\,\delta f + \sqrt{\frac{m}{G}}, \tag{11}$$

which, assuming $G(\delta f)^2 \gg 1$, is maximized for

$$m_{\rm opt} = \frac{1}{16G(\delta f)^2}, \tag{12}$$

at which point,

$$\left(\frac{df}{dt}\right)_{\rm opt} = \frac{1}{8G(\delta f)}, \tag{13}$$

which means that the rate of increase of fitness $F = fG$ is

$$\frac{dF}{dt} = \frac{1}{8(\delta f)} \text{ per generation.} \tag{14}$$

For a population with low fitness ($\delta f < 0.125$), the rate of increase of fitness may exceed 1 unit per generation. Indeed, if $\delta f \lesssim 1/\sqrt{G}$, the rate of increase, if $m = 1/2$, is of order $\sqrt{G}$; this initial spurt can only last of order $\sqrt{G}$ generations. For $\delta f > 0.125$, the rate of increase of fitness is smaller than one per generation. As the fitness approaches G, the optimal mutation rate tends to $m = 1/(4G)$, so that an average of $1/4$ bits are flipped per genotype, and the rate of increase of fitness is also equal to $1/4$; information is gained at a rate of about 0.5 bits per generation. It takes about $2G$ generations for the genotypes of all individuals in the population to attain perfection.

For fixed m, the fitness is given by

$$\delta f(t) = \frac{1}{2\sqrt{mG}}(1 - ce^{-2mt}), \tag{15}$$

subject to the constraint $\delta f(t) \leq 1/2$, where c is a constant of integration, equal to 1 if $f(0) = 1/2$. If the mean number of bits flipped per genotype, mG, exceeds 1, then the fitness F approaches an equilibrium value $F_{\rm eqm} = (1/2 + 1/(2\sqrt{mG}))G$.

If m is tuned to the optimal fitness–dependent value, $m_{\rm opt}$ (12) then the fitness is given, assuming $\delta f(0) = 0$, by

$$\delta f(t) = \frac{t^{1/2}}{2\sqrt{G}}, \tag{16}$$

which hits $\delta f = 1/2$ at $t = G$.

This theory is somewhat inaccurate in that the true probability distribution of fitness is non–Gaussian, asymmetrical, and quantized to integer values. All the same, the predictions of the theory are not grossly at variance with the results of simulations in section 2.4.

2.2 Theory of sex

If we assume an infinitely large population then we do not need to keep track of the individual genotypes, because descendants of one individual are unlikely to marry. We assume furthermore that the population is in a homogeneous state $f_g = f(t)$, for all g.

How does $f(t+1)$ depend on $f(t)$? Let's first assume the two parents of a child both have exactly $f(t)G$ good bits, and, by our homogeneity assumption, that those bits are independent random subsets of the G bits. The number of bits that are good in both parents is roughly $f(t)^2 G$, and the number that are good in one parent only is roughly $2f(t)(1-f(t))G$, so the fitness of the child will be $f(t)^2 G$ plus the sum of $2f(t)(1-f(t))G$ fair coin flips, which has a binomial distribution of mean $f(t)(1-f(t))G$ and variance $\frac{1}{2}f(t)(1-f(t))G$. The fitness of a child is thus roughly distributed as

$$F_{\rm child} \sim \text{Normal}\left(\text{mean}=f(t)G, \text{variance}=\frac{1}{2}f(t)(1-f(t))G\right). \tag{17}$$

The important property of this distribution, contrasted with the distribution under mutation, is that the mean fitness is equal to the parents' fitness; the variation produced by sex does not reduce the average fitness.

We now include the parental population's variance, which we will write as $\sigma^2(t) = \beta(t)\frac{1}{2}f(t)(1-f(t))G$. The average of the two parents will have variance $\sigma^2(t)/2$, so the population of all children will have fitness, before selection, distributed as

$$F_{\rm child} \sim \text{Normal}\left(\text{mean}=f(t)G, \text{variance}=\left(1+\frac{\beta}{2}\right)\frac{1}{2}f(t)(1-f(t))G\right). \tag{18}$$

Natural selection selects the children on the upper side of this distribution. The mean increase in fitness will be

$$\bar{F}(t+1) - \bar{F}(t) = [\alpha(1+\beta/2)^{1/2}/\sqrt{2}]\sqrt{f(t)(1-f(t))G}, \tag{19}$$

and the variance of the surviving children will be

$$\sigma^2(t+1) = \gamma(1+\beta/2)\frac{1}{2}f(t)(1-f(t))G, \tag{20}$$

where $\alpha = \sqrt{2/\pi}$ and $\gamma = (1-2/\pi)$. If there is dynamic equilibrium $[\sigma^2(t+1) = \sigma^2(t)]$ then

$$\gamma(1+\beta/2) = \beta, \text{ so } (1+\beta/2) = \frac{2}{2-\gamma}, \tag{21}$$

and the factor in (19) is

$$\alpha(1+\beta/2)^{1/2}/\sqrt{2} = \sqrt{\frac{2}{(\pi+2)}} \simeq 0.62. \tag{22}$$

Defining this constant to be $\eta \equiv \sqrt{2/(\pi+2)}$, we conclude that, under sex and natural selection, the mean fitness of the population increases at a rate *proportional to the square root of the size of the genome*,

$$\frac{d\bar{F}}{dt} \simeq \eta\sqrt{f(t)(1-f(t))G} \text{ bits per generation.} \tag{23}$$

If, recklessly, we take our homogeneity assumption to hold for all time, we can write $\bar{F} = fG$ and obtain the differential equation:

$$\frac{df}{dt} \simeq \frac{\eta}{\sqrt{G}}\sqrt{f(t)(1-f(t))}, \tag{24}$$

whose solution is

$$f(t) = \frac{1}{2}\left[1+\sin\left(\frac{\eta}{\sqrt{G}}(t+c)\right)\right], \quad \text{for } t+c \in \left(-\tfrac{\pi}{2}\sqrt{G}/\eta, \tfrac{\pi}{2}\sqrt{G}/\eta\right), \tag{25}$$

where c is a constant of integration, $c = \sin^{-1}(2f(0)-1)$. So this idealized system reaches a state of eugenic perfection ($f = 1$) within a finite time: $(\pi/\eta)\sqrt{G}$ generations.

2.3 Finite populations

In a population of size N, this analysis becomes inaccurate: after only $\log N$ generations, most married couples share ancestors, and their genotypes are thus not independent. We can still make the weak statement that, for the first $\log N$ generations, information will be acquired at a rate proportional to $\sqrt{G}$. However, it seems that sex mixes the genes of the population sufficiently rapidly that the infinite population approximation works well for many more generations, as long as the population N is not too small. For small enough N, we expect that, whilst the average fitness of the population increases, some unlucky bits may become frozen into the bad state. (These bad genes are sometimes known as hitchhikers.) Eventually, we expect the population to reach a homogeneous state in which all individuals have identical genotypes that are mainly 1–bits, but have some 0–bits too. The smaller the population, the greater the number of frozen 0–bits is expected to be.

Figure 3 shows results as a function of population size N with $G = 100$ and $G = 1000$. The infinite population approximation appears to work well for remarkably small population sizes. Only when N is smaller than $\sim \sqrt{G}$ do the empirical results deviate strongly from the theoretical curve. If N is significantly smaller than $\sqrt{G}$, information cannot possibly be acquired at a rate as big as $\sqrt{G}$, since the information content of the blind watchmaker's decisions cannot be any greater than $2N$ bits per generation, this being the number of bits required to specify which of the $2N$ children get to reproduce. Baum *et al.* (1995), analyzing a similar model, show that the population size N should be about $\sqrt{G}(\log G)^2$ to make hitchhikers unlikely to arise.

2.4 Simulations

Figure 2(a) shows the fitness of a population of $N = 1000$ individuals with a genome size of $G = 1000$ starting from a random initial state with normalized fitness 0.5. It also shows the

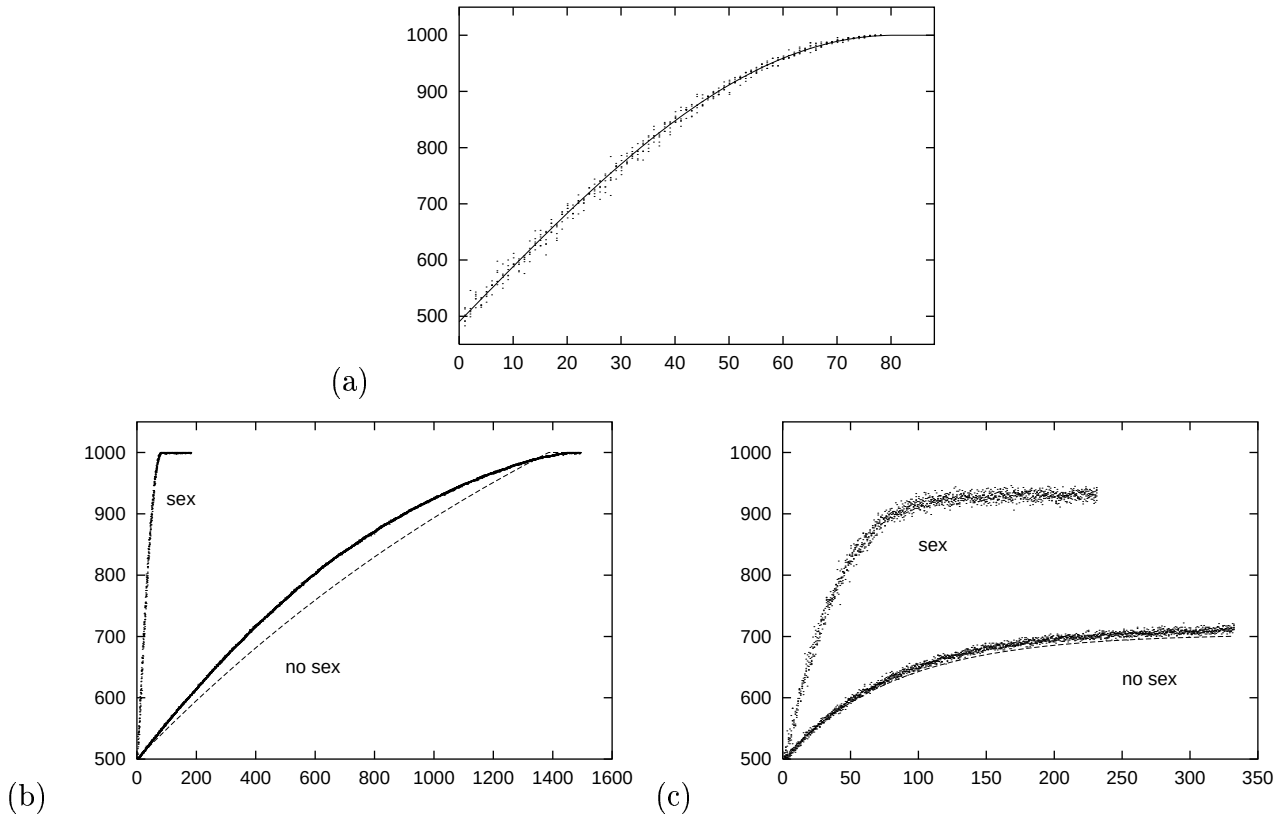

Figure 2: Fitness as a function of time. The genome size is $G = 1000$.
The dots show the fitness of six randomly selected individuals from the birth population at each generation. The initial population of $N = 1000$ had random generated genomes with $f(0) = 0.5$ (exactly). (a) Variation produced by sex alone. Line shows theoretical curve (25) for infinite homogeneous population.
(b) Variation produced by mutation, with and without sex, when the mutation rate is $mG = 0.25$ bits per genome. The dashed line shows the curve equation (15).
(c) Variation produced by mutation, with and without sex, when the mutation rate is $mG = 6$ bits per genome.

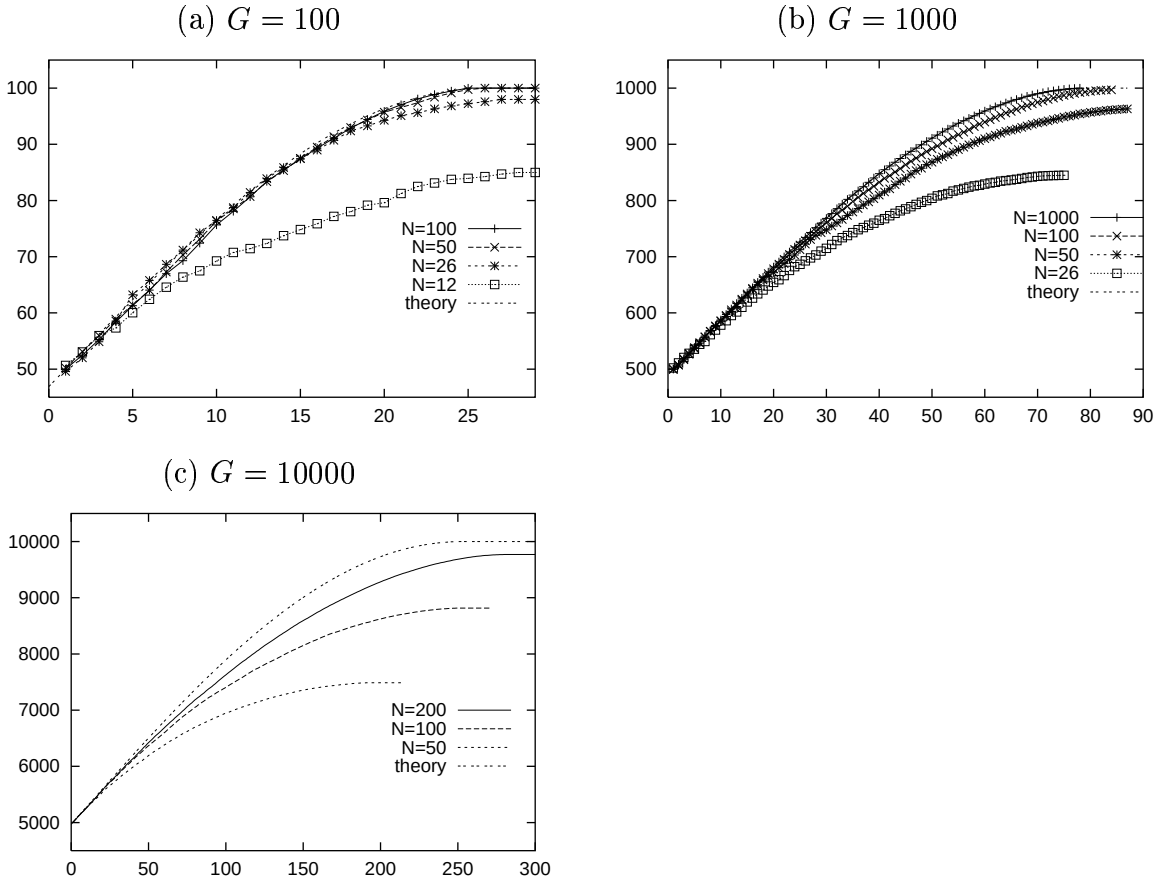

Figure 3: Dependence of results on population size for a sexual population with no mutations. Each graph is the mean fitness of the birth population.

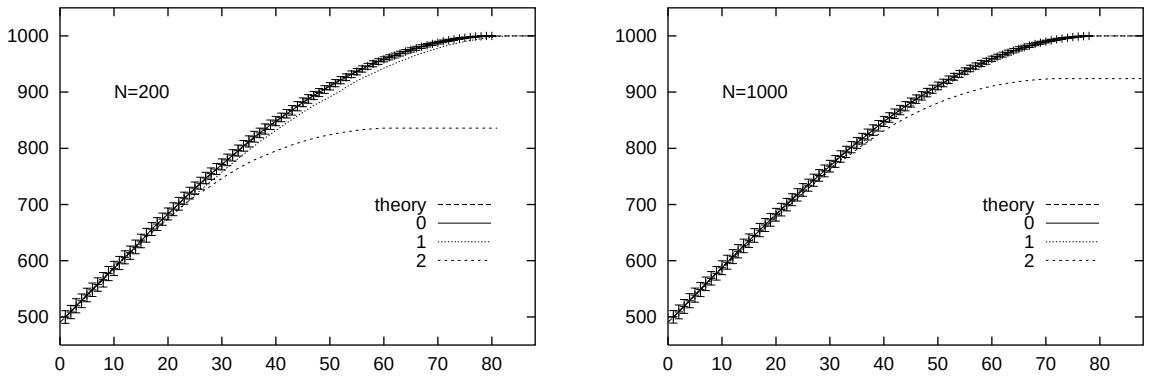

Figure 4: Dependence of results on crossover mechanism. Each graph is the mean fitness of the birth population. For method 0, the standard deviation is also indicated by error bars.

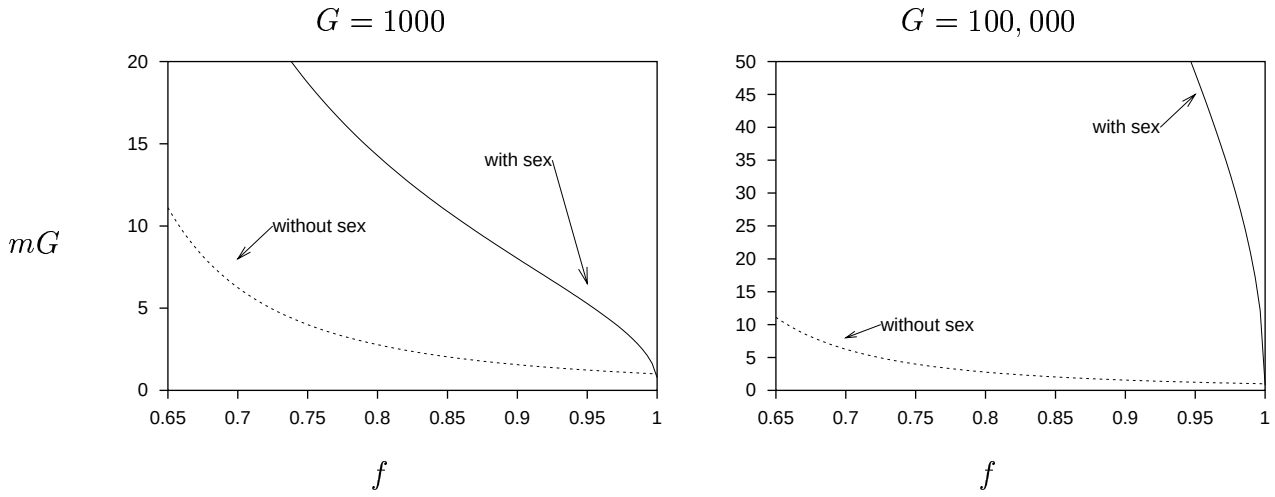

Figure 5: Maximal tolerable mutation rate, shown as number of errors per genome (mG), versus normalized fitness $f = F/G$. Independent of genome size, a parthenogenetic species can only tolerate of order 1 error per genome per generation; a species that uses recombination can tolerate far greater mutation rates.

theoretical curve $f(t)G$ using $f(t)$ derived for the infinite homogeneous population, equation (25), which fits remarkably well.

In contrast, figures 2(b1,2) show the evolving fitness when variation is produced by mutation at rates $0.25/G$ and $6/G$ respectively. Note the difference in the horizontal scales.

3 The maximal tolerable mutation rate

This section needs checking over, to confirm the details of the factors of η, etc.

What if we combine the two models of variation? What is the maximum mutation rate that can be tolerated by a species that has sex?

If a child, produced by sex between parents whose excess normalized fitness is δf, suffers mutations with probability m per bit, then the child's δf will have a normal distribution with mean $(1-2m)\delta f$ and variance

$$\frac{1}{G}\left[m + \frac{1}{2}f(1-f)\right], \tag{26}$$

where we have assumed $m \ll 1$, as before. The rate of increase of fitness is given by[1]

$$\frac{df}{dt} \simeq -2m\,\delta f + \eta\sqrt{2}\sqrt{\frac{m + f(1-f)/2}{G}}, \tag{27}$$

which is positive if

$$2m\,\delta f < \eta\sqrt{2}\sqrt{\frac{m + f(1-f)/2}{G}}. \tag{28}$$

[1] Aside: the *optimal* mutation rate in this model (to maximize df/dt) is zero, if $\delta f > 0$.

Replacing δf by its largest value, $1/2$, and omitting the m on the right hand side, the rate of increase of fitness is positive, for a given f, if the mutation rate satisfies

$$m < \eta \sqrt{\frac{f(1-f)}{G}}. \tag{29}$$

Let us compare this rate with the result in the absence of sex, which, from equation (11), is that the maximum tolerable mutation rate is

$$m < \frac{1}{G} \frac{1}{(2\,\delta f)^2}. \tag{30}$$

These two maximum mutation rates are of completely different orders. The tolerable mutation rate with sex is of order $\sqrt{G}$ times greater than that without sex!

A parthenogenetic (non–sexual) species could try to wriggle out of this bound on its mutuation rate by increasing its litter sizes. But if mutation flips on average mG bits, the probability that no bits are flipped in one genome is roughly e^{-mG}, so a mother needs to have roughly e^{mG} offspring in order to have a good chance of having one child with the same fitness as her. The litter size of a non–sexual species thus has to be exponential in mG, if mG is bigger than 1, if the species is to persist.

So the maximum tolerable mutation rate is pinned close to $1/G$, for a non–sexual species, whereas it is a larger number of order $1/\sqrt{G}$, for a species with recombination.[2]

3.1 Dependence on the model of crossover

The original crossover method (method 0) set each child bit at random either to the corresponding bit of parent 1 or parent 2. We have also simulated two other models of crossover, both of which model the one–dimensional local topology of a chromosome. In method 1, there is a crossover probability χ before each bit. [The special case $\chi = 0.5$ gives method 0.] In method 2, crossovers only occur at hotspots located every d bits along the chromosome; the probability of crossover at a hotspot is here set to 0.5.

Figure 4 shows results for the three methods for various population sizes. For method 1, $\chi = 0.05$, and for method 2, $d = 10$, so these two methods have the same mean number of crossovers. We find that, for large populations, the results for methods 1 and 2 are similar to those for method 0. However, the population size below which freezing of bad genes is seen is larger for method 1 and larger still for method 2.

[2]Turning these results around, we can predict the largest possible genome size for a given fixed mutation rate, m. For a parthenogenetic species, the largest genome size is of order $1/m$, and for a sexual species, $1/m^2$. Taking the figure $m = 10^{-8}$ as the mutation rate per nucleotide per generation (Eyre-Walker and Keightley, 1999), and allowing for a maximum brood size of $20,000$ (that is, $mG \simeq 10$), we predict that all species with more than 10^9 coding nucleotides make at least occasional use of recombination. If the brood size is 12, then this number falls to 2.5×10^8. This crude calculation used mutation–rate figures from humans and apes, so let's do it again for prokaryots. If the mutation rate is $m = 10^{-7}$ (need a reference) per base pair per generation for a prokaryot, and $mG \leq 1$, the maximum genome size is about 10^7 nucleotides. What then, of the Bdelloid rotifers, the fresh–water invertebrates that appear to be entirely parthenogenetic? If the Bdelloid genome is similar in size to *C. Elegans* — 6×10^7 coding basepairs, let's say — then this creature falls below the rough threshold, 2.5×10^8, derived above. My guess is that the organism evolved rapidly using sex, until it fitted its niche perfectly; thereafter, its genome size was small enough that it had no need for sex to keep itself healthy. I wait to learn the genome size.

4 Discussion

These results quantify the well known argument for why species reproduce by sex with recombination, namely that recombination allows useful mutations to spread more rapidly through the species and allows deleterious mutations to be more rapidly cleared from the population (Maynard Smith, 1978; Felsenstein, 1985; Maynard Smith, 1988; Maynard Smith and Száthmary, 1995). A population that reproduces by recombination can acquire information from natural selection at a rate of order $\sqrt{G}$ times faster than a parthenogenetic population, and it can tolerate a mutation rate that is of order $\sqrt{G}$ times greater. For genomes of size $G \simeq 10^8$ coding nucleotides, this factor of $\sqrt{G}$ is substantial.

This enormous advantage conferred by sex has been noted before by Kondrashov (1988), but this meme, which Kondrashov calls 'the deterministic mutation hypothesis', does not seem to have diffused throughout the evolutionary research community, as there are still numerous papers in which the prevalence of sex is viewed as a mystery to be explained by elaborate mechanisms (Kelling and Rand, 1995).

Maynard Smith and Száthmary (1995) assert that deleterious mutations are eliminated more effectively under recombination only if the effects of mutations have a synergistic rather than an antagonistic effect on fitness; they then argue that mutations in metabolic enzymes are likely to act antagonistically in unicellular organisms and synergistically in higher organisms, thus accounting for the general lack of sex in the former and its prevalence in the latter. I think this explanation is overly complicated. Simple organisms can do without sex because they have small genomes.

Of course, this conclusion depends on the fitness function that we assume, and on our model of selection. Is it reasonable to model fitness, to first order, as a *sum* of independent terms? Maynard Smith (1968) argues that it is: the more good genes you have, the higher you come in the pecking order, for example. The directional selection model has been used extensively in theoretical population genetic studies (Bulmer, 1985). We might expect real fitness functions to involve interactions, in which case crossover might reduce the average fitness. However, since recombination gives the biggest advantage to species whose fitness functions are additive, we might predict that evolution will have favoured species that used a representation of the genome that corresponds to a fitness function that has only weak interactions. And even if there are interactions between genes, it seems plausible that the fitness would still involve a sum of such interacting terms, with the number of terms being some fraction of G. Furthermore, if the fitness function were a highly nonlinear function of the genotype, it could be made more smooth and locally linear by the Baldwin effect. The Baldwin effect (Baldwin, 1896; Hinton and Nowlan, 1987) has been widely studied as a mechanism whereby *learning* guides evolution, and it could also act at the level of transcription and translation. Consider the evolution of a peptide sequence for a new purpose, assuming the effectiveness of the peptide is highly nonlinear function of the sequence, perhaps having a small island of good sequences surrounded by an ocean of equally bad sequences. In an organism whose transcription and translation machinery is flawless, the fitness will be an equally nonlinear function of the DNA sequence and progress towards the island will be by a random walk. In contrast, an organism having the same DNA sequence, but whose DNA–to–RNA transcription or RNA–to–protein translation is 'faulty', will occasionally, by mistranslation or mistranscription, produce a working enzyme, and will do so with greater probability if its DNA sequence is close to a good sequence. One cell might produce 1000 proteins from the one mRNA sequence, of which 999 have no enzymatic effect, and one does. The one working catalyst will be enough for that cell to have an increased fitness relative to rivals whose DNA sequence is further from the island of good sequences. For this reason we conjecture that, at least early in evolution, and perhaps still now,

the genetic code was not implemented perfectly but was implemented noisily, with some codons coding for a distribution of possible amino acids. This noisy code could even be switched on and off from cell to cell in an organism by having multiple aminoacyl–tRNA synthetases, some more reliable than others.

Is it reasonable to use truncation selection, which kills off all children with fitness below a threshold? (In using this selection model, we are not implying that those children were unfit to reproduce — merely that they were the losers in the competition for limited resources.) Other selection functions could be used; any monotonic selection function, concave or convex, is expected to give similar results, since the advantage of recombination is that it gives a variance in fitness proportional to G, without decreasing the mean fitness. The results obtained with 'ranking selection', in which pairs of random individuals are compared and the fitter is selected, are equivalent to the results derived here for threshold selection, except that the rate of evolution is smaller by a factor of $\sqrt{2}$.

The rate at which information can be acquired by natural selection was addressed by Kimura (1961); using two different arguments, he asserts that the maximum rate of information acquisition is 1 bit per generation (assuming that a generation is the time taken for the population to double). That the 'speed limit for evolution' is about one bit per generation is also 'proved' by Worden (1995), who goes beyond Kimura by showing that this result holds whether or not the species has sex. This contrasting conclusion is reached because of Worden's choice of model. Instead of defining an individual's fitness as a relative quantity, involving competition with other individuals, his model assumes that one's genotype determines one's probability of having children absolutely.

Whilst our model assumed that the bits of the genome do not interact, ignored the fact that the information is represented redundantly, assumed that there is a direct relationship between phenotypic fitness and the genotype, and assumed that the crossover probability in recombination is high, we believe these qualitative results would still hold if more complex models of fitness and crossover were used: the relative benefit of sex will still scale as $\sqrt{G}$. Only in small, in–bred populations are the benefits of sex expected to be diminished.

Stability of a gene for sex or parthenogenesis

At Ewan Birney's suggestion, I modified the model so that one of the G bits in the genome determines whether an individual prefers to reproduce parthenogenetically ($x = 1$) or sexually ($x = 0$). The big advantage of parthenogenesis, from the point of view of the individual, is that one is able to pass on 100% of one's genome to one's children, instead of only 50%. The results depend on the number of children had by a single parthenogenetic mother, $K_{\rm p}$ and the number of children born by a sexual couple, $K_{\rm s}$. Both ($K_{\rm p} = 2$, $K_{\rm s} = 4$) and ($K_{\rm p} = 4$, $K_{\rm s} = 4$) are reasonable models. The former ($K_{\rm p} = 2$, $K_{\rm s} = 4$) would seem most appropriate in the case of unicellular organisms, where the cytoplasm of both parents goes into the children. The latter ($K_{\rm p} = 4$, $K_{\rm s} = 4$) is appropriate if the children are solely nurtured by one of the parents, so single mothers have just as many offspring as a sexual pair. I concentrate on the latter model, since it gives the greatest advantage to the parthenogens, who are supposedly expected to outbreed the sexual community. Because parthenogens have four children per generation, the maximum tolerable mutation rate for them is twice the expression derived before for $K_{\rm p} = 2$ (30). If the fitness is large, the maximum tolerable rate is $mG \simeq 2$.

Initially the genomes are set randomly with $F = G/2$, so half of the population have the gene for parthenogenesis. Figure 6 shows the outcome if single parthenogens produce as many offspring as a *pair* of sexuals. During the 'learning' phase of evolution, in which the fitness is increasing rapidly, pockets of parthenogens appear briefly, but then disappear within a couple of generations as their

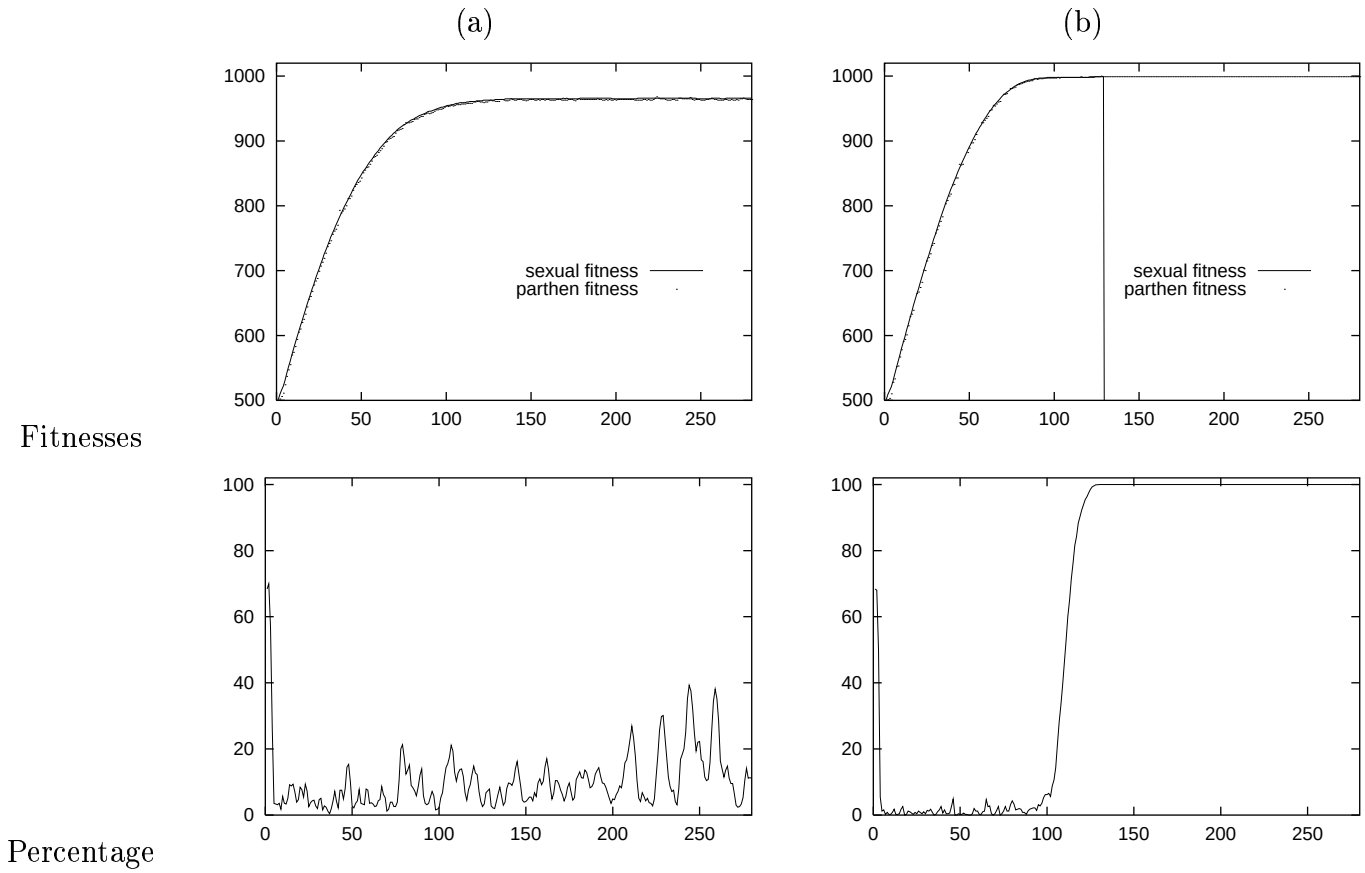

Figure 6: Results when there is a gene for parthenogenesis, and no interbreeding, *and single mothers produce as many children as sexual couples.* $G = 1000$, $N = 1000$. (a) $mG = 4$; (b) $mG = 1$. Vertical axis shows both fitness and percentage of the population that is parthenogenetic.

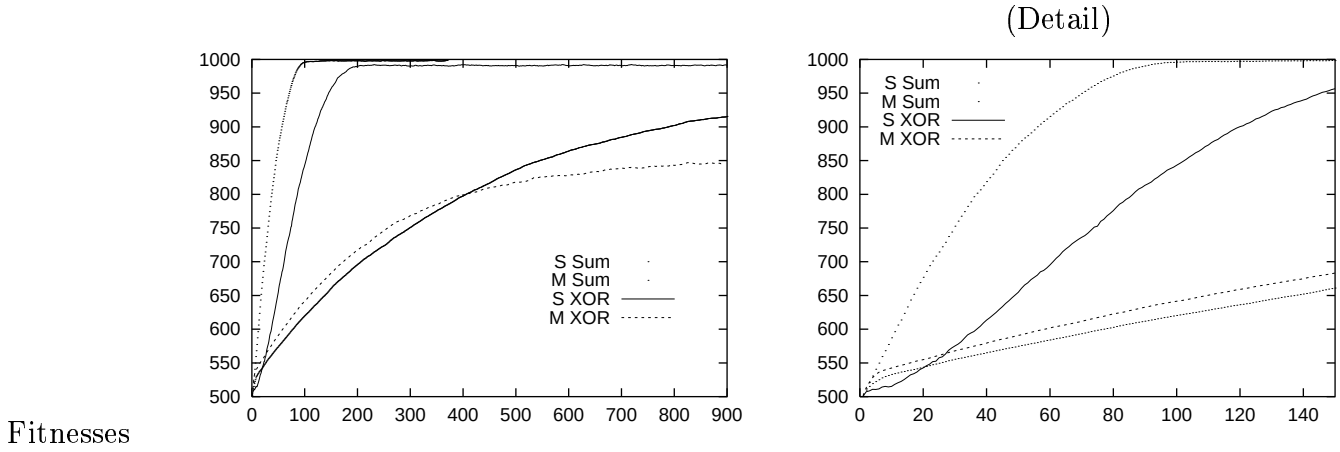

Figure 7: Results for a fitness function with interactions. The graphs show the fitnesses as a function of time for four separate populations with and without sex, and with an interacting and a non–interacting fitness function. These four populations did not compete with each other in these simulations.

sexual cousins overtake them in fitness and leave them behind. Once the population reaches its top fitness, however, the parthenogens can take over, if the mutation rate is sufficiently low ($mG = 1$). In these simulations, sex does not tend to reappear once the parthenogens have taken over, because a small sexual community, having size $N_{\rm sexual} < \sqrt{G}$, will be in–bred and will not have the advantage discussed in the rest of this paper.

In the presence of a higher mutation rate ($mG = 4$), however, the parthonegens never take over. The breadth of the sexual population's fitness is of order $\sqrt{G}$, so a mutant parthenogenetic colony arising with slightly above–average fitness will last for about $\sqrt{G}/(mG) = 1/(m\sqrt{G})$ generations before its fitness falls below that of its sexual cousins. As long as the population size is sufficiently large for some sexual individuals to survive for this time, sex will not die out.

In a sufficiently unstable environment, where the fitness function is continually changing, the parthenogens will always lag behind the sexual community. These results are consistent with the argument of Haldane and Hamilton that sex is helpful in an arms race with parasites. The parasites define an effective fitness function which changes with time, and a sexual population will always ascend the current fitness function more rapidly.

5 Interactions

What happens if we modify the fitness function such that there are interactions between bits in the genome? A simple such fitness function is

$$F_{\rm XOR}(\mathbf{x}) = \sum_{g=1}^{G/2} x_g {\rm XOR} x_{g+G/2}, \tag{31}$$

which can be contrasted with

$$F_{\rm noninteracting}(\mathbf{x}) = \sum_{g=1}^{G/2} x_g, \tag{32}$$

which has the same statistics, that is, any value of fitness can be realised by just the same number of vectors $\mathbf{x}$ for either function.

Figure 7 shows the four outcomes for sexual and asexual populations evolving under these two fitness functions with $G = 2000$ and $N = 100$; the mutation rate is set to $mG = 2$. These experiments allow us to study two questions. First, if there are interactions among bits, does sex still have an advantage over sexlessness? Second, if there are two competing species, one that has a fitness function that does not have interactions, and one that does, which species will dominate?

The answer to the first question is shown by the heavy solid and dashed lines: early on in evolution, sex has a disadvantage, because crossover splits up linked bits. However, the sexual population eventually becomes fitter than the asexual one because it is better able to eliminate mutations.

The answer to the second question is that the species with the non–interacting fitness function becomes fitter faster. Thus there is an evolutionary pressure in favour of genomes that are organized to have a simple fitness function.

Acknowledgements

This work was stimulated by a talk given by John Maynard Smith. I thank Sanjoy Mahajan, Robert MacKay, Ken Miller, Ewan Birney, Ted Briscoe, Oscar Gaggiotti, Jonathan Swinton, Geoff Hinton and Graeme Mitchison for helpful discussions.

After completion of this work my attention was drawn to a similar paper by Alex Rogers and Adam Prügel–Bennett, and to publications on similar models of an evolving sexual population in the Genetic Algorithms community (Mühlenbein and Schlierkamp-Voosen, 1993; Baum *et al.*, 1995); these papers, which note that fitness increases at a rate proportional to $\sqrt{G}$, concern the relative advantages of various genetic algorithms, all of which use sex.

References

Baldwin, J. (1896) A new factor in evolution. *American Naturalist* **30**: 441–451.

Baum, E., Boneh, D., and Garrett, C. (1995) On genetic algorithms. In *Proceedings of the Eighth Annual Conference on Computational Learning Theory*, pp. 230–239, New York. ACM.

Bulmer, M. (1985) *The Mathematical Theory of Quantitative Genetics*. Oxford: Oxford University Press.

Eyre-Walker, A., and Keightley, P. (1999) High genomic deleterious mutation rates in hominids. *Nature* **397**: 344–347.

Felsenstein, J. (1985) Recombination and sex: is Maynard Smith necessary? In *Evolution. Essays in honour of John Maynard Smith*, ed. by P. J. Greenwood, P. H. Harvey, and M. Slatkin, pp. 209–220. Cambridge: C.U.P.

Hinton, G., and Nowlan, S. (1987) How learning can guide evolution. *Complex Systems* **1**: 495–502.

Kelling, M. J., and Rand, D. A. (1995) A spatial mechanism for the evolution and maintenance of sexual reproduction. *Oikos* **74**: 414–424.

Kimura, M. (1961) Natural selection as the process of accumulating genetic information in adaptive evolution. *Genetical Research Cambridge*.

Kondrashov, A. S. (1988) Deleterious mutations and the evolution of sexual reproduction. *Nature* **336** (6198): 435–440.

Maynard Smith, J. (1968) "Haldane's dilemma" and the rate of evolution. *Nature* **219** (5159): 1114–1116.

Maynard Smith, J. (1978) *The Evolution of Sex*. Cambridge: C.U.P.

Maynard Smith, J. (1988) *Games, Sex and Evolution*. Hertfordshire: Harvester–Wheatsheaf.

Maynard Smith, J., and Száthmary, E. (1995) *The Major Transitions in Evolution*. Oxford: Freeman.

Mühlenbein, H., and Schlierkamp-Voosen, D. (1993) Predictive models for the breeder genetic algorithm I. continuous parameter optimization. *Evolutionary Computation* **1**: 25–50.

Worden, R. P. (1995) A speed limit for evolution. *Journal of Theoretical Biology* **176** (1): 137–152.