Multidimensional epistasis and the disadvantage of sex

Fyodor A. Kondrashov* and Alexey S. Kondrashov

National Center for Biotechnology Information, National Institutes of Health, Building 38A, Bethesda, MD 20892

Edited by John Maynard Smith, University of Sussex, Brighton, United Kingdom, and approved July 31, 2001 (received for review May 1, 2001)

Sex is thought to facilitate accumulation of initially rare beneficial mutations by allowing simultaneous allele replacements at many loci. However, this advantage of sex depends on a restrictive assumption that the fitness of a genotype is determined by fitness potential, a single intermediate variable to which all loci contribute additively, so that new alleles can accumulate in any order. Individual-based simulations of sexual and asexual populations reveal that under generic selection, sex often retards adaptive evolution. When new alleles are beneficial only if they accumulate in a prescribed order, a sexual population may evolve two or more times slower than an asexual population because only asexual reproduction allows some overlap of successive allele replacements. Many other fitness surfaces lead to an even greater disadvantage of sex. Thus, either sex exists in spite of its impact on the rate of adaptive allele replacements, or natural fitness surfaces have rather specific properties, at least at the scale of intrapopulation genetic variability.

Fisher (1) and Muller (2) noticed that new, beneficial mutations may be incorporated into a sexual population faster than into an asexual population if recombination creates initially absent genotypes that carry multiple mutant alleles. This effect depends on a finite population size (3, 4). However, sex also can facilitate adaptive evolution regardless of random drift, as long as fitness grows at less than an exponential rate with the number of beneficial alleles in the genotype, which must be the case as fitness cannot approach infinity (5, 6). Thus, sex is thought to speed up allele replacements driven by directional selection under a broad range of plausible conditions, although the importance of this effect for the evolution of sex remains controversial (7–12).

The analysis that led to this conclusion was performed under the assumption that fitness is a function of a single, additive variable, fitness potential (13, 14). The fitness potential of a genotype can be the genotype's breeding value of a quantitative trait or the sum of contributions from all of the beneficial mutations, deleterious mutations, or heterozygous loci. Unless fitness depends exponentially on fitness potential, selection defined by such a function is called epistatic (6).

However, even an arbitrary, epistatic fitness surface defined by a function of a fitness potential still represents a rare exception within the space of all possible fitness surfaces, because a truly generic fitness surface is not a function of any single, additive genotype-determined variable. In the case of just two loci with impacts x and y, a fitness surface f(x,y) usually cannot be represented as $f(x,y) = F(p)$, where $p = x + y$ is a fitness potential. In particular, multiple fitness peaks not arranged on the same straight line or curved ridges of high fitness do not fit into the fitness-potential Procrustean bedstead. We will call a generic fitness surface multidimensionally epistatic to distinguish it from unidimensionally epistatic fitness functions of a fitness potential. First considered by Bateson (15), Dobzhansky (16), and Muller (17), multidimensional epistasis was studied mostly in the context of speciation (18–21) and some related phenomena (22–24). Here, we investigate its impact on the evolution of sex.

N-Dimensional Epistasis. Let us start from the case opposite to that of fitness potential and assume that at some L loci new beneficial alleles 1 can replace old alleles 0 only in a prescribed order. Without loss of generality, we assume that fitness is an increasing function of the number of loci at the beginning of a genotype that are all occupied by alleles 1, i.e., that genotypes 000. . . , 100. . . , 110. . . , and 111. . . have increasingly high fitnesses, whereas alleles 1 that follow one or several alleles 0 (e.g., in 011. . .) are mildly deleterious. Thus, to evolve from the initial genotype 000. . . 0 into the best genotype 111. . . 1, a population must follow the only possible succession of allele replacements. Geometrically, this assumption means that the fitness surface contains a single narrow ridge of increasing fitness connecting 000. . . 0 and 111. . . 1 that turns into a new direction perpendicular to all its previous directions every time it reaches a corner of the L-dimensional hypercube of all possible genotypes. In contrast, fitness that depends on fitness potential is an increasing function of the total number of alleles 1 in the genotype regardless of their order, so that the total number of possible passes within the space of genotypes from 000. . . 0 to 111. . . 1 is L!.

We used an individual-based model with unidirectional ($0 \rightarrow 1$) mutation. Sexual reproduction consisted of random mating followed by free recombination. The fitness of an individual was $(1 + a)^n \times (1 - d)^m$, where n is the number of loci at the beginning of the genotype that are all occupied by alleles 1, m is a number of alleles 1 preceded by allele(s) 0, a is the advantage of an allele 1 not preceded by allele(s) 0, and d is the disadvantage of an allele 1 preceded by allele(s) 0. Simulations were performed using METROWERKS C program, which is available on request.

This extreme form of multidimensional epistasis leads to a substantial disadvantage of sex (Fig. 1). In a sexual population, successive allele replacements proceed with little overlap, because allele 1 at locus k becomes advantageous only after allele 1 reaches a high frequency at locus $k - 1$, since otherwise recombination would produce too many maladapted genotypes . . . 01. . . . In contrast, successive allele replacements overlap substantially in an asexual population because the genotype . . . 11. . . , once formed by mutation, multiplies regardless of the rest of the population. One can say that the extreme form of multidimensional epistasis causes synergistic epistasis between beneficial alleles 1 at the successive loci (because $w_{\ldots 00\ldots} \times w_{\ldots 11\ldots} > w_{\ldots 01\ldots} \times w_{\ldots 10\ldots}$, where $w_{\ldots ij\ldots}$ is the average fitness of individuals having allele i at locus $k - 1$ and allele j at locus k), which generally favors asexual reproduction (3, 5). The disadvantage of sex disappears only when the mutation rate becomes lower than $1/N$, forcing nonoverlapping allele replacements even in an asexual population (data not reported). Thus, the impact of sex in this case is the opposite to that in the case of fitness potential, because the transition from fitness potential-based selection to the extreme form of multidimensional epistasis slows down the evolution of a sexual population much more than the evolution of an asexual population.

This paper was submitted directly (Track II) to the PNAS office.

*To whom reprint requests should be addressed. E-mail: fkondras@ncbi.nlm.nih.gov.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

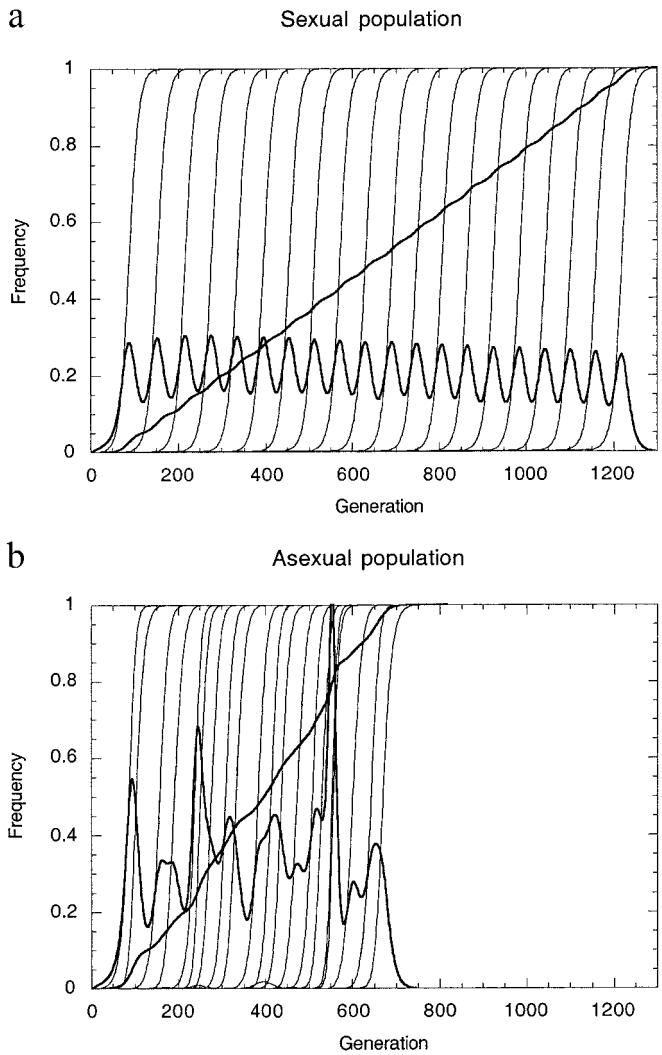

Fig. 1. Accumulation of new beneficial mutations by sexual (*a*) and asexual (*b*) populations in the case of extreme multidimensional epistasis. The number of loci L is 20, the population size N is 10^6, the per locus mutation rate μ is 3×10^{-5}, the advantage of an allele 1 a is 0.1, and the disadvantage of an allele 1 preceded by 0 d is 0.01. The average frequency of allele 1 at all of the loci (bold line, increasing), the variance in the number of alleles 1 per genotype (bold line, fluctuating), and the frequencies of allele 1 at individual loci (thin lines) are shown.

Two-Dimensional Epistasis. Biologically, a strictly prescribed order of accumulation of new beneficial alleles implied by the extreme case of multidimensional epistasis may be too extreme: perhaps mutations at more than one locus are beneficial at any moment. However, an arbitrary order implied by the fitness potential model also appears to be an oversimplification: it is impossible to perfect fingers before hands come into existence. Thus, an intermediate model is worth considering. Let us assume that all of the loci are split into two sets of loci, L_1 and L_2 ($L_1 + L_2 = L$), and fitness of a genotype is determined by two subfitness potentials, the total numbers of alleles 1 within L_1 and L_2. Geometrically, fitness will now be a function of two variables instead of one variable (fitness potential) or L variables (extreme multidimensional epistasis). We will assume that this function forms the narrowest possible continuous ridge connecting the points (0,0) and (L_1,L_2). Such a ridge in any discrete multidimensional genotype space can consist only of straight lines and straight corners—otherwise, it would either be wide or discontinuous (Fig. 2). The fitness of genotypes belonging to the ridge

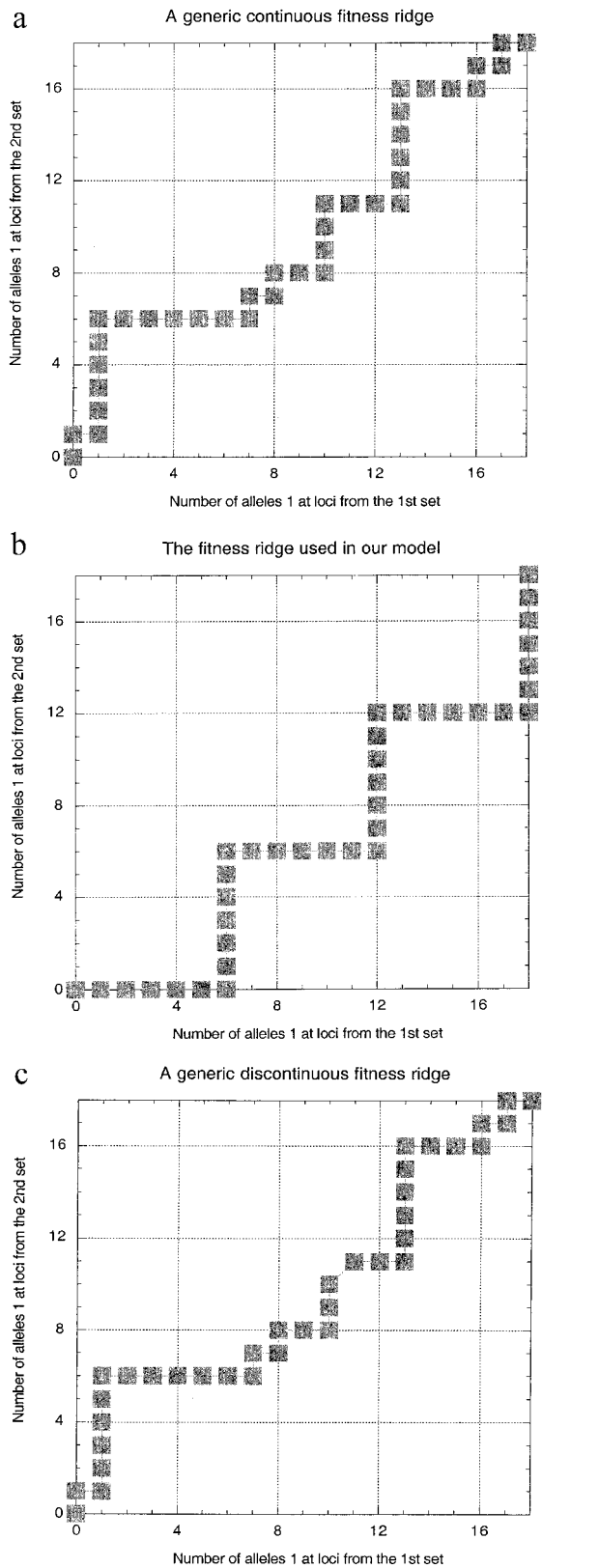

Fig. 2. (*a*) An example of a generic, continuous, narrow fitness ridge connecting the points (0,0) and (18,18) where a point (p,q) corresponds to any genotype that has p alleles 1 at the first L_1 loci and q alleles 1 at the rest L_2 loci ($L_1 = L_2 = 18$). (*b*) A continuous fitness ridge used in our simulations. (*c*) A discontinuous fitness ridge such that, to evolve from the phenotype (10,10) to the phenotype (11,11), a population must acquire two alleles 1 that are individually deleterious.

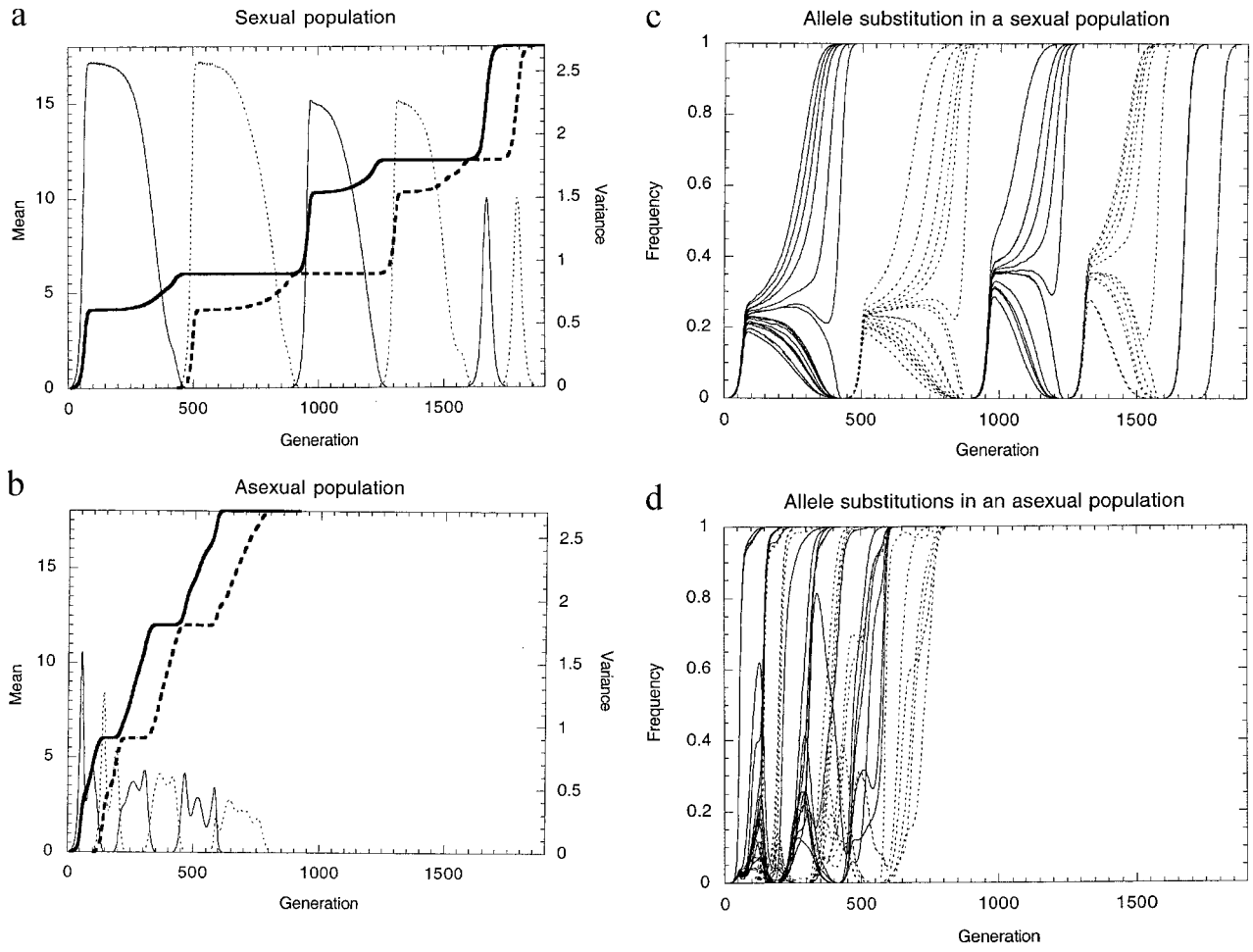

Fig. 3. Accumulation of new beneficial mutations by sexual (*a*, *c*) and asexual (*b*, *d*) populations in the case of a continuous, narrow fitness ridge, as shown in Fig. 2*b*. The population size N is 10^6, the per locus mutation rate μ is 3×10^{-5}, the advantage of an allele 1 a is 0.1, and the rate of fitness decline caused by deviation of a genotype from fitness ridge δ is 0.3. The means (thick lines) and variances (thin lines) of the numbers of allele 1 at the loci from the first group (solid lines) and from the second group (broken lines) are shown in *a* and *c*. The frequencies of allele 1 at individual loci from the first group (solid lines) and from the second group (broken lines) are shown in *b* and *d*.

of high fitness was $(1 + a)^n$, where n is the total number of alleles 1 in a genotype and a is the advantage of an allele 1 for such genotypes. The fitness of genotype G outside the fitness ridge was $(1 + a)^n \times (1 - \delta)^D$, where δ is the disadvantage of deviating from the fitness ridge, D is the Hemming distance of G from the closest genotype within the fitness ridge, and n is the number of alleles 1 within this closest genotype (if several genotypes within the ridge were equidistant from G, the genotype with the highest n was chosen).

A sexual population evolves much faster than an asexual population along straight regions of the continuous fitness ridge (Fig. 2*b*) but slows down drastically as it approaches a corner (Fig. 3). This deceleration is caused by the following sequence of events. For a sexual population to turn the first corner of the fitness ridge, i.e., to begin accumulating allele 1 at the second group of 18 loci, almost all individuals in the population must carry exactly six alleles 1 at the first group of 18 loci. Because of recombination, this condition requires near fixations of alleles 1 at some 6 loci from the first group, and near fixations of alleles 0 at the remaining 12 loci. However, when the population approaches the corner, frequencies of allele 1 initially are almost uniform across all of the first 18 loci. Then, selection acting on these loci becomes essentially stabilizing, causing magnification of initially small, random differences in allele frequencies across these loci (25) and eventual fixations of the desired numbers of alleles 0 and 1. Under $N\mu > 1$, these differences are small, leading to very slow allele fixations. Thus, paradoxically, the rate of evolution of a sexual population increases when the number of beneficial mutants per locus per generation declines. In contrast, an asexual population does not slow down when it approaches a corner, because alleles 1 at the second group of loci can accumulate in any clone that acquired six alleles 1 at the first group of loci. As a result, a sexual population evolves much slower than an asexual population (Fig. 3), unless straight regions of a fitness ridge are very long. We cannot easily explain these results in terms of pairwise epistatic interactions among the loci.

The disadvantage of sex is even more drastic when a fitness ridge is discontinuous (Fig. 2*c*). An asexual population can cross a discontinuity (fitness valley) because of fixation of a double mutant . . . 11. . . , whereas a sexual population cannot do so because a rare double mutant will mate with the original genotype . . . 00. . . , after which recombination will produce maladapted genotypes . . . 01. . . and . . . 10. . . (5, 26, 27).

Discussion

To summarize, it seems that unless selection can be approximated by the fitness potential model, sexual reproduction usually impedes, rather than facilitates, fixations of new, beneficial alleles. Our results seem to be robust and hold under a variety of values of a, d, and δ (data not reported). A random fitness

surface that includes continuous fitness ridges connecting genotypes 00. . . 0 and 11. . . 1 contains a lot of corners and does not fit into the fitness potential model (20, 24). Of course, natural fitness surfaces may be very different from those generated randomly.

Indirect evidence suggests that one of a million pairs of independently acquired new mutations, if present within the same genotype, turn out to be strongly incompatible and cause drastically reduced fitness (28), thus implying that the characteristic length of a straight region of the fitness ridge is ≈1,000. However, this estimate is rather imprecise and does not take into account the possibility of ecologically mediated multidimensional epistasis. Still, it is possible that fitness potential is a good approximation at a relatively small scale of intrapopulation variability (making it applicable, for example, to selection against deleterious mutations) but cannot be used at a larger scale of interspecific differences. Thus, analysis of fitness in multidimensional genotype spaces is an extremely important experimental problem. Unless data will show that ridges of high fitness are mostly straight and rarely contain corners, facilitation of adaptive evolution cannot be the reason for the origin and maintenance of sex.

1. Fisher, R. A. (1930) *The Genetical Theory of Natural Selection* (Clarendon, Oxford).
2. Muller, H. J. (1932) *Am. Nat.* **66,** 118–138.
3. Crow, J. F. & Kimura, M. (1965) *Am. Nat.* **99,** 439–450.
4. Felsenstein, J. (1974) *Genetics* **78,** 737–756.
5. Eshel, I. & Feldman, M. W. (1970) *Theor. Popul. Biol.* **1,** 88–100.
6. Shnol, E. E. & Kondrashov, A. S. (1993) *Genetics* **134,** 995–996.
7. Maynard Smith, J. (1978) *The Evolution of Sex* (Cambridge Univ. Press, Cambridge, U.K.).
8. Maynard Smith, J. (1980) *Genet. Res.* **35,** 269–277.
9. Maynard Smith, J. (1988) *Genet. Res.* **51,** 59–63.
10. Kondrashov, A. S. (1993) *J. Hered*. **84,** 372–387.
11. Hurst, L. D. & Peck, J. R. (1996) *Tree* **11,** 46–52.
12. Barton, N. H. & Charlesworth, B. (1998) *Science* **281,** 1986–1990.
13. Milkman, R. (1978) *Genetics* **88,** 391–403.
14. Kimura, M. & Crow, J. F. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 6168–6171.
15. Bateson, W. (1909) in *Darwin and Modern Science*, ed. Seward, A. S. (Cambridge Univ. Press, Cambridge, U.K.), pp. 85–101.
16. Dobzhansky, T. (1936) *Genetics* **21,** 113–135.
17. Muller, H. J. (1942) *Biol. Symp*. **6,** 71–125.
18. Kondrashov, A. S. & Mina, M. V. (1986) *Biol. J. Linn. Soc.* **27,** 201–223.
19. Orr, H. A. (1996) *Genetics* **144,** 1331–1335.
20. Gavrilets, S. (1999) *Am. Nat*. **154,** 1–22.
21. Gavrilets, S., Li, H. & Vose, M. D. (1998) *Proc. R. Soc. London Ser. B* **265,** 1483–1489.
22. Gavrilets, S. (1997) *Evolution (Lawrence, Kans.)* **51,** 1027–1035.
23. Gavrilets, S. (1999) *Proc. R. Soc. London Ser. B* **266,** 817–824.
24. Gavrilets, S. & Gravner, J. (1997) *J. Theor. Biol.* **184,** 51–64.
25. Barton, N. H. (1989) *Genet. Res.* **54,** 59–77.
26. Michalakis, Y. & Slatkin, M. (1996) *Genet. Res.* **67,** 257–269.
27. Phillips, P. C. (1996) *Genet. Res.* **67,** 271–283.
28. Orr, H. A. (1995) *Genetics* **139,** 1805–1813.