Genet. Res., Camb. (1982), **40**, *pp.* 325–332
With 3 text-figures
Printed in Great Britain

SHORT PAPER

Selection against harmful mutations in large sexual and asexual populations

By ALEXEY S. KONDRASHOV

Research Computer Centre of the Academy of Sciences of the USSR, Pushchino, Moscow Region, 142292, USSR

(*Received 4 December 1981 and in revised form 21 April 1982*)

SUMMARY

Selection against harmful mutations in large populations is studied assuming that the rate of fitness decrease grows with every new mutation added to a genome. Under this reasonable assumption (Mayr, 1970) the average fitness of a sexual population, without linkage between the loci, is higher, and the average equilibrium number of harmful mutations per individual lower, than in an asexual population. If a gamete contains on the average one or more new mutations, the resulting advantage of sexual reproduction and recombination seems to be sufficient to counterbalance the double advantage of parthenogenesis. Moreover, selection against harmful mutations is probably the most powerful factor preventing linkage disequilibrium even with epistatic interaction between the loci.

1. INTRODUCTION

The efficacies of selection against harmful mutations in sexual and asexual populations were first compared by Muller (1964). He noted that, while in a finite asexual population elimination of all individuals free of harmful mutations is irreversible (Muller's ratchet), such individuals may reappear in a sexual population due to recombination. This accounts for the fact that sexual reproduction may be advantageous in small populations. This advantage becomes large when the size of the population and the selection against single mutations are sufficiently small and the mutation rate is high (Maynard Smith, 1978, ch. 3).

For infinite populations, Maynard Smith concludes that the type of reproduction has no influence on selection against harmful mutations (Maynard Smith, 1978, ch. 3). However, he considers only multiplicative selection, which provides an exponential decrease of individuals' fitness with an increase of mutation number in the genome. Besides, it implies that the difference between the fitness of individuals with i and $i+1$ mutations is less than that of individuals with $i-1$ and i mutations. Though different types of selection against harmful mutations are poorly investigated, many authors do consider such an assumption unreasonable. Mayr (1970, ch. 9) considers that the decrease in fitness of an individual with three harmful genes, each of which has 1 % lower fitness than the normal allele, will be much more than 3 %, reaching perhaps 10 %. This and other similar assumptions mean that every harmful mutation that follows the last leads to a

https://doi.org/10.1017/S0016672300019194 Published online by Cambridge University Press

larger decrease of fitness, so that only individuals having harmful mutations less than a certain critical value (k) are viable. The limiting cases are, on the one hand the linear selection under which every additional mutation leads to the same decrease of fitness and, on the other hand, the threshold selection under which individuals with less than k mutations all have the same fitness and the rest are inviable (Fig. 1). The efficacy of threshold selection against harmful mutations has been recently shown (Kimura & Crow, 1979). In the present paper the efficacy of selection of this type will be compared in infinite sexual and asexual populations using a somewhat different approach. Two models will be considered: the first one is simplified to illustrate the idea of this paper, the second presents more detail.

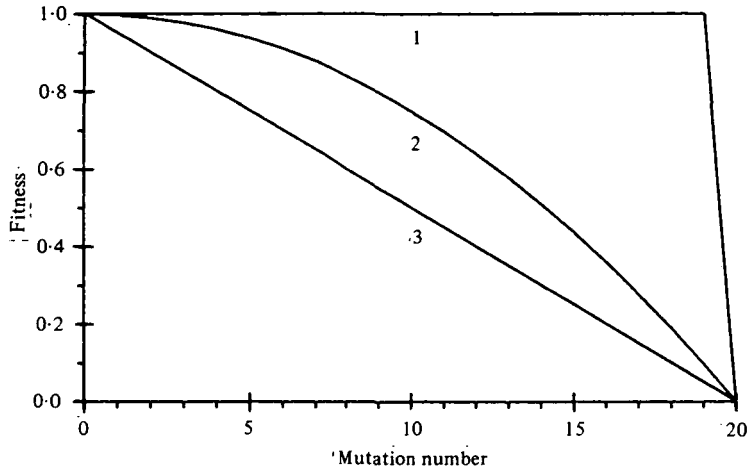

Fig. 1. Fitness of phenotypes with different mutation numbers under threshold (1), intermediate (2) and linear (3) selection with $k = 20$.

Since mutations may be regarded as harmful if they cannot be unrestrictedly accumulated in genomes of the individuals, we shall in all cases consider stationary states of populations when accumulation of the mutations is balanced by the work of selection. Assume that there is no linkage between the loci and the mutant allele is rare at each locus so that our models may be applied to haploids as well as to diploids. In the latter case mutations occur only in heterozygotes. Mutations at all loci will be assumed equally harmful, so that an individual's fitness is determined by the number of mutations in its genome.

According to Mukai (Mukai *et al.* 1972) in populations of *Drosophila melanogaster* the average number of new harmful mutations per genome (u) is not less than 0·9 for each generation. It is possible, though, that this value is underestimated by a factor of 2 or even more because of underestimation of the number of minor effect mutations (Kimura & Crow, 1979). Organisms with large genomes and a large number of replications per generation are likely to have an even higher rate of mutation. According to Mukai an equilibrium number of harmful mutations per genome is approximately 20–50 (Kimura & Crow, 1979). These data underlie the parameter values in our models.

(i) *Model 1*

Let threshold selection act against harmful mutations. Then all individuals in an asexual population at equilibrium will have exactly $k-1$ mutations after selection, and

https://doi.org/10.1017/S0016672300019194 Published online by Cambridge University Press

since mutations occur independently at all loci, their progenies will be viable with probability $l_a = e^{-u}$, i.e. under the condition that an offspring contains no new mutations.

Consider now a sexual population whose life cycle is: mutations–recombination–selection–mutations. Let us assume, following Maynard Smith, that alleles at all the loci are independently distributed after recombination: then the frequency of individuals with i mutations in their genomes will be

$$q_i = \frac{e^{-\lambda}\lambda^i}{i!}$$

where λ is the average number of mutations per genome after recombination. The average number of mutations per genome after selection is

$$\mu = \sum_{i=0}^{k-1} \frac{\lambda^i}{(i-1)!} \cdot \left(\sum_{i=0}^{k-1} \frac{\lambda^i}{i!}\right)^{-1}.$$

Table 1. *Model 1 investigation results*

(*a*) Average mutation number per individual in a sexual population (μ)

u/k	5	20	80
0·5	2·271	13·06	62·96
2·0	3·010	15·22	67·81
8·0	3·604	17·26	72·95

(*b*) Average fitness of a sexual population

u/k	5	20	80
0·5	0·8521	0·9399	0·9748
2·0	0·4387	0·7180	0·8758
8·0	0·0100	0·1230	0·4431

(*c*) Average fitness of an asexual population

$u = 0·5$	0·6065
$u = 2·0$	0·1353
$u = 8·0$	0·00034

(*d*) Advantage of a sexual reproduction

u/k	5	20	80
0·5	1·405	1·550	1·6072
2·0	3·2413	5·305	6·471
8·0	29·82	366·8	1320·0

Then in an equilibrium population $\lambda - \mu = u$. It is easy to solve this equation on a computer. The probability of survival of the progeny in a sexual population is

$$l_s = e^{-\lambda} \sum_{i=0}^{k-1} \frac{\lambda^i}{i!},$$

and the advantage of sexual reproduction $-l_s l_a^{-1}$. The results are given in Table 1.

(ii) *Model 2*

Let the fitness of an individual having i harmful mutations be $s_i = 1 - i/k^\alpha$. Then in the case $\alpha = 1$ the selection is linear and with $\alpha \to \infty$ it is threshold. The case $\alpha = 2$, which will be called intermediate selection, is also considered here (Fig. 1).

Let the frequency of individuals of a certain generation with i mutations in their genome be p_i in an asexual population after selection. Then, as an offspring has j new mutations with probability $e^{-u}(u^j/j!)$, the frequency of individuals with i mutations in their genome is

$$p'_i = e^{-u} \sum_{j=0}^{i} p_i \frac{u^{i-j}}{(i-j)!} \tag{1}$$

after the mutation process.

https://doi.org/10.1017/S0016672300019194 Published online by Cambridge University Press

The distribution after selection becomes:

$$P_i = p'_i \cdot s_i \cdot l_a^{-1} \tag{2}$$

where the average fitness of progeny is

$$l_a = \sum_{i=0}^{k-1} p'_i \cdot s_i. \tag{3}$$

In a sexual population, the frequencies of individuals in a given generation with i mutations before and after the mutation process, q_i and q'_i, are connected by an equation similar to (1).

To describe the recombination process more accurately than in the previous model note that since mutations at each locus are rare, the frequency with which an offspring from

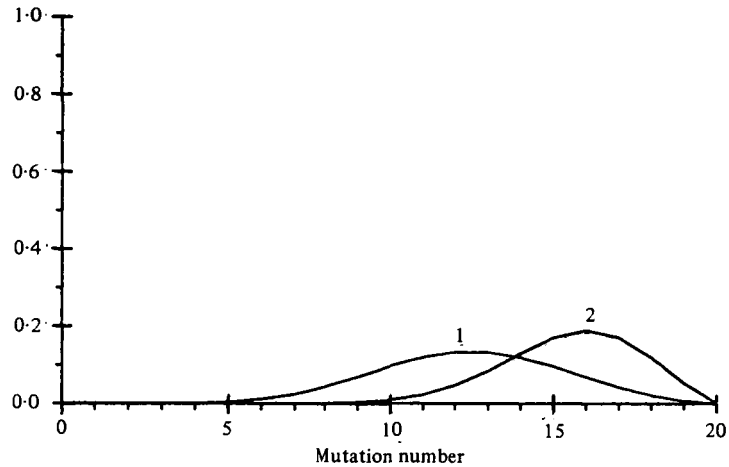

Fig. 2. Distributions of mutation numbers in genome in equilibrium sexual (1) and asexual (2) populations under Fig. 1 linear selection, $u = 2$.

parents with n and m mutations has i mutations is $b_{n,m,i} = C^i_{n+m}(\frac{1}{2})^i \cdot (\frac{1}{2})^{n+m-i}$. Therefore, after recombination, the distribution of individuals is connected with q'_i by the formula: $q''_i = \Sigma_{n+m \geqslant i} q'_n \cdot q'_m \cdot b_{n,m,i}$. The distribution of individuals after selection may be deduced from q''_i using

$$Q_i = q''_i \cdot s_i \cdot l_s^{-1}; \quad l_s = \sum_{i=0}^{k-1} q''_i \cdot s_i \tag{4}$$

similar to (2). The average fitness of progeny in a sexual population l_s is determined in the same way as for an asexual one, and the selective value of sexual reproduction is $l_s \cdot l_a^{-1}$. Stationary values of the average number of mutations per individual, the fitness of sexual and asexual populations, and the selective value of sexual reproduction were calculated on a computer. To achieve sufficient accuracy, from 20 to 500 generations were needed depending on parameter values and the initial choice of distribution. It was found that with any type of selection the average fitness of a stationary asexual population is e^{-u} (Table 1C). The author is obliged to Yu. G. Zarhin for the following proof of the fact. Consider the first non-zero element of a stationary distribution $\tilde{p}_h$. Then $\tilde{p}'_h = e^{-u}\tilde{p}_h$ and as $\tilde{P}_i = \tilde{p}_i$, $l_a = e^{-u}$. Note that under linear and intermediate selection even $\tilde{p}_o \neq 0$, while under threshold selection – $\tilde{p}_{k-1} \neq 0$. Other computation results are presented in Fig. 2, Table 2.

https://doi.org/10.1017/S0016672300019194 Published online by Cambridge University Press

We also investigated the process of competition between sexual and asexual reproduction in the same population with asexual individuals having a twofold advantage. To this effect formulas

$$P_i = 2p'_i s_i \cdot \left(\sum_{i=0}^{k-1} 2p'_i s_i + q''_i \cdot s_i \right)^{-1} \tag{2'}$$

and

$$Q_i = q''_i \cdot s_i \cdot \left(\sum_{i=0}^{k-1} 2p'_i \cdot s_i + q''_i \cdot s_i \right)^{-1} \tag{4'}$$

Table 2. *Model 2 Investigation Results*

		(*a*)				(*b*)	
u/k	5	20	80	*u/k*	5	20	80
0·5	4·000	19·00	79·00	0·5	0·8323	0·9324	0·9724
	2·521	13·78	64·31		1·372	1·537	1·603
2·0	4·000	19·00	79·00	2·0	0·3799	0·6618	0·8519
	3·202	16·04	69·78		2·8074	4·8802	6·2950
8·0	4·000	19·00	79·00	8·0	0·005812	0·06674	0·3273
	3·671	17·74	74·75		17·33	193·9	975·6
		(*c*)				(*d*)	
0·5	1·473	7·369	30·98	0·5	0·6453	0·6577	0·6655
	1·282	6·313	26·26		1·064	1·084	1·097
2·0	2·952	15·47	69·59	2·0	0·2132	0·2846	0·3184
	2·478	12·40	52·53		1·575	2·103	2·353
8·0	3·709	18·34	78·24	8·0	0·002171	0·01510	0·05507
	3·365	16·44	69·50		6·468	44·96	164·2
		(*e*)				(*f*)	
0·5	2·453	11·75	48·51	0·5	0·7285	0·7800	0·7948
	1·758	8·442	35·25		1·201	1·2860	1·3105
2·0	3·415	16·88	73·05	2·0	0·2799	0·4106	0·4735
	2·722	13·16	55·60		2·0684	3·0341	3·4989
8·0	3·833	18·54	78·31	8·0	0·003314	0·02541	0·09692
	3·449	16·59	69·90		9·8600	75·76	288·9

(*a*), (*c*), (*e*). Average mutation number in equilibrium asexual (the top figure) and sexual (the lower one) populations.

(*b*), (*d*), (*f*). Sexual population fitness (top) and advantage of sexual reproduction (bottom).

(*a*), (*b*). Threshold; (*c*), (*d*), linear; (*e*), (*f*), intermediate selection.

were substituted for (2) and (4). Here Q_i and P_i signify the parts of the whole population which makes up sexual and asexual individuals, respectively, with i mutations in their genome,

$$\Sigma_{i=0}^{k-1} Q_i \text{ and } \Sigma_{i=0}^{k-1} P_i$$

are the total frequencies of all sexual and asexual individuals, respectively. Figure 3 shows a typical fate of an asexual subpopulation under such conditions which make the advantage of sexual reproduction greater than 2. Otherwise, sexual individuals were outlived (data not presented).

3. DISCUSSION

It follows from Tables 1 and 2 that under all investigated conditions a balanced sexual population contains a smaller mutation average than an asexual one and has a better fitness. The advantage of sexual reproduction increases along with the growth of rate

https://doi.org/10.1017/S0016672300019194 Published online by Cambridge University Press

of mutation and k, i.e. with the decrease of individual effect of single mutations. It may be explained by the fact that with the large equilibrium number of mutations in a sexual population, death of each individual leads to elimination of a larger number of mutations, and thus fitness of the population increases. In a number of cases the advantage of sexual reproduction approximates the maximum of e^{-u} which is reached with the fitness of the sexual population equal to 1. The advantage of sexual reproduction is maximal under the threshold selection and is minimal under the linear one. Our supposition (in model 1) about independent distribution of alleles after recombination leads to a marked overestimation of the advantage of sexual reproduction only at high mutation frequency.

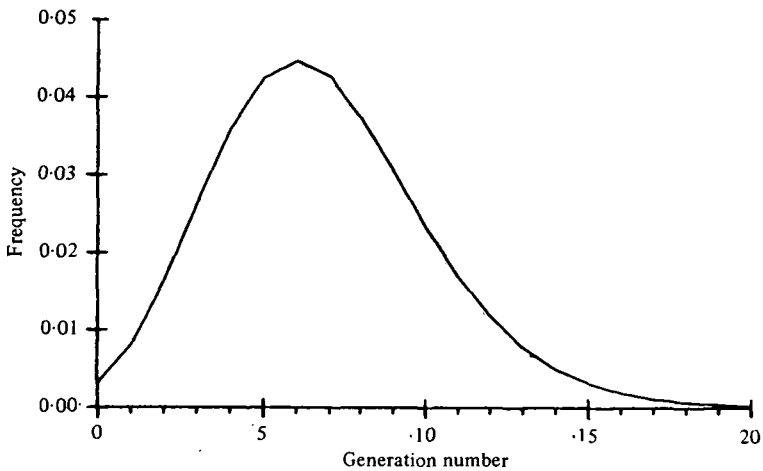

Fig. 3. Frequency of asexual clone in a sexual population in successive generations under Fig. 1 linear selection, $u = 2$. Initial number of mutations in asexual individuals is 10, sexual population being in equilibrium.

Recently, the question about the nature of the factors outbalancing the twofold advantage of parthenogenesis (Maynard Smith, 1978, ch. 1) has been widely discussed. It is well known that sexual reproduction is advantageous if the population evolves directionally (see review by Maynard Smith, 1978, ch. 2), but this mechanism cannot be likely to maintain sexual reproduction in slowly evolving groups which have not changed essentially in tens of millions of years.

It seems that sexual reproduction may become advantageous under unpredictable changes of environment (Maynard Smith, 1980). It is difficult to suppose, however, that all sexual populations are influenced by constant environmental fluctuations which are sufficient to compensate for the parthenogenetic advantage. That is why the mechanisms noted above, as well as Muller's ratchet working only in small populations, can hardly be responsible for the wide occurrence of sexual reproduction.

Our results (Table 2) suggest that if the mutation rate is more than two mutations per genome in each generation then sexual reproduction has more than a twofold advantage. Due to the universal nature of the mutation process, this makes it possible to suppose that it is selection against harmful mutations that is the main factor maintaining sexual reproduction, though more consistent data on mutation rate as well as on the types of selection against harmful mutations are lacking. Under some other types of selection the difference between sexual and asexual populations stands out even more remarkably. For example, consider a population where individuals having more than the average number of mutations in their genomes die and all the rest have equal

https://doi.org/10.1017/S0016672300019194 Published online by Cambridge University Press

fitness. Then, if the population is asexual, mutations unlimitedly accumulate there, and if it is sexual their number tends to a limit depending on the mutation rate.

It follows from our data that the advantage of sexual reproduction with regard to group selection determines the fate of asexual clones in a sexual population, which is in line with Maynard Smith's conclusion that selection against asexual individuals is always group selection (Maynard Smith, 1978, the last sentence in the book).

If selection against harmful mutations does play a role in maintaining sexual reproduction, then in asexual and parthenogenetic populations the mutation rate must be much less than in sexual ones, which is in good accord with experimental data (Leslie & Vrijenhoek, 1978), though an explanation involving Muller's ratchet is also possible. High frequency of polyploids among parthenogenetic forms can also be explained by inefficiency of selection against harmful mutations when some of them are partially recessive (E. Kh. Ginzburg, personal communication). Though in the long run polyploidy cannot decrease the mutation load, it can provide a relatively long period of existence for a parthenogenetic clone, after its formation.

As sexual reproduction is ineffective without recombination, selection against harmful mutations must lead to increase in recombination frequency (Feldman *et al.* 1980). It seems that recombination frequency in wild life is a compromise between selection for recombination and for linkage. If something causes a decrease of selection for linkage, then before a new recombination frequency is fixed those gametes which have been formed with larger than average crossover number will have the advantage. This may account for the fact that the decrease of linkage disequilibrium in cage *Drosophila* populations goes faster than it would do without selection under the same recombination rate (Clegg *et al.* 1980).

The latest data on DNA sequences indicate that elimination of harmful mutations is a regular function of selection (Li *et al.* 1981). It confirms our hypothesis about the advantages of sexual reproduction and explains the absence of systematic deviations from linkage equilibrium in natural populations (see, e.g. Loukas *et al.* 1981) because selection against mutations 'mixes up' the genome notwithstanding possible epistatic interactions between loci.

The author is obliged to S. E. Shnol for drawing his attention to this problem, to A. D. Bazykin and A. S. Rozanov for helpful discussions, and to Natasha Kondrashova for translation.

REFERENCES

Clegg, M. T., Kidwell, J. F. & Horch, C. R. (1980). Dynamics of correlated genetic systems. V. Rates of decay of linkage disequilibria in experimental population of *Drosophila melanogaster*. *Genetics* **94**, 217–234.

Feldman, M. W., Christiansen, F. B. & Brooks, L. D. (1980). Evolution of recombination in constant environment. *Proceedings of the National Academy of Sciences of the USA* **77**, 4838–4841.

Kimura, M. & Crow, J. F. (1979). Efficiency of truncation selection. *Proceedings of the National Academy of Sciences of the USA* **76**, 396–399.

Leslie, J. F. & Vrijenhoek, R. C. (1978). Genetic dissection of clonally inherited genomes of *Poecilopsis*. I. Linkage analysis and preliminary assessment of deleterious gene loads. *Genetics* **90**, 801–811.

Li, W.-H., Gojobori, T. & Nei, M. (1981). Pseudogenes as a paradigm of neutral evolution. *Nature* **292**, 8237–8239.

Loukas, M., Vegrini, Y. & Krimbas, C. B. (1981). The genetics of *Drosophila subobscura* populations. XVII. Further genic heterogeneity within electromorphs by urea denaturation and the effect of the increased genic variability on linkage disequilibrium studies. *Genetics* **97**, 429–441.

https://doi.org/10.1017/S0016672300019194 Published online by Cambridge University Press

MAYNARD SMITH, J. (1978). *The Evolution of Sex*. Cambridge University Press.
MAYNARD SMITH, J. (1980). Selection for recombination in a polygenic model. *Genetical Research* **35**, 269–278.
MAYR, E. (1970). *Populations, species and evolution*. Harvard University Press, Cambridge, Mass.
MUKAI, T., CHIGUSA, S. T., METTLER, L. E. & CROW, J. F. (1972). Mutation rate and dominance of genes affecting viability in *Drosophila melanogaster*. *Genetics* **72**, 335–355.
MULLER, H. J. (1964). The relation of recombination to mutational advance. *Mutation Research* **1**, 2–9.

https://doi.org/10.1017/S0016672300019194 Published online by Cambridge University Press