bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Sex Solves Haldane's Dilemma.

Donal A. Hickey
Department of Biology, Concordia University
7141 Sherbrooke West,
Montreal, Quebec H4B 1R6
Canada

and

G. Brian Golding
Department of Biology, McMaster University
1280 Main St. West,
Hamilton, Ontario L8S 4K1
Canada

bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

Haldane's dilemma refers to the fact that selection acting at many loci could incur an unsustainable reproductive cost. This is because natural selection involves the removal of disfavored genotypes from the population and their replacement with selectively favored genotypes. Consequently, it requires a certain amount of reproductive excess to maintain a constant population size. The required reproductive excess is referred to as the cost of selection. It is generally assumed that this cost is cumulative when several genetic loci are subject to independent selection. Consequently, even if the amount of reproductive excess is sufficient to allow allelic replacement at a single locus, it has been argued that it may not be sufficient to allow many genes to undergo independent selective replacement simultaneously. This cumulative cost of multi-locus selection has been seen as a potentially limiting factor on the rate of adaptive evolution. In this paper we show that Haldane's arguments for the accumulation of costs are valid for a clonally reproducing population of asexual genotypes. But recombination, in a randomly mating sexual population, eliminates the accumulation of costs over multiple loci. Thus, sex removes a perceived reproductive constraint on the rate of adaptive evolution. These results are particularly relevant to recent evolutionary models of adaptation based on the standing genetic variation.

bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Introduction

Natural selection depends on reproductive excess because disfavored alleles must be culled in order for favored alleles to increase in frequency within the selected population. Such culling, in the absence of a sufficient reproductive excess, would result in a decrease in population size and, eventually, it could lead to the extinction of the population. Haldane (1957, 1960) referred to the culling of disfavored types as the "cost of selection". In addition to defining this cost, Haldane estimated how great this cost might be during the substitution of one allele for another at a single locus. He then argued that, given independent selection at different genetic loci, this cost would be approximately additive over loci (Haldane 1957, 1960). In other words, if the cost of allelic substitution at one locus equals x, then the cost for L loci would equal Lx. This would mean that evolutionary adaptations that involved allelic substitutions at many genetic loci could prove to be "too costly", given the available reproductive excess. Thus, simultaneous selection at many loci might lead to extinction of the population or, at best, it would constrain the rate of adaptive change. This problem has been referred to as Haldane's dilemma (VanValen 1963).

Haldane's calculations of the cost of natural selection had a major impact on the field of evolutionary biology. For example, Dodson (1962) wondered if the cost were too high to allow the adaptive divergence between humans and chimpanzees, while Kimura (1968, 1995) used the concept of selective cost to support his theory of neutral evolution (see Kern and Hahn 2018 for a recent review). Other studies have pointed out various possible ways in which the cost of allele substitution at many loci might be reduced. But all of these solutions involve changing some of the starting assumptions. For example, Maynard Smith (1968) and Sved (1968) pointed out that the cost could be reduced if we assume some form of truncation selection rather than independent selection. Grant and Flake (1974) showed that the cost could be reduced in a sub-divided population. These solutions are not very intuitively satisfying, however, since they are contingent on specific forms of selection or population structure. Moreover, while these modifying assumptions can reduce the accumulation of costs over loci, they do not eliminate it. The crucial point, however, is that Haldane (1957) did not take recombination into account. In this paper, we use numerical simulations to demonstrate that Haldane's argument about

bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the summation of the costs over multiple loci does not apply to a sexually recombining population – and this is true regardless of the selection scheme or the population structure. In the Discussion below, we explain why Haldane's concept of cumulative costs applies only to a quasi-infinite asexual population. A randomly outbreeding sexual population, on the other hand, avoids the problem envisaged by Haldane.

Haldane's reasoning (Haldane 1957) for the cost of selection contains two parts. First, there is the calculation of the reproductive cost of substituting one allele for another at a single locus. Secondly, there is the summation of these costs over many loci. In general, the time to fixation is a function of the reproductive capacity and the initial frequency (see Felsenstein, 1971). The second part of Haldane's argument is based on the fact that the expected frequency of genotypes carrying multiple favorable alleles (at different loci, given linkage equilibrium between loci) is lower than the frequency of genotypes carrying a single favorable allele. Thus, it would take approximately twice as long for a genotype carrying two favorable alleles to reach fixation as it would for a single mutant (Grant and Flake 1974). As we demonstrate in this study, it is this accumulation of costs over several loci that is avoided through recombination in a sexually outbreeding population. Specifically, in a sexual population we do not need to assume an initial population that is large enough to contain at least one individual with the optimum genotype. And it is precisely this assumption of a very large initial population that led Haldane to his conclusion that the costs of selection would accumulate over loci.

In fairness to Haldane, it should be noted that his examples of selection at multiple loci referred explicitly to populations with asexual reproduction. For instance, he wrote the following: *Consider clonally reproducing bacteria, in which a number of disadvantageous genes are present, kept in being by mutation, each with frequencies of* 10^{-4}*. They become slightly advantageous through a change of environment or residual genotype. Among* 10^{12} *bacteria, there might be one which possessed three such mutants. But since the cost of selection is proportional to the negative logarithm of the initial frequency the mean cost of selecting its descendants would be the same as that of selection for the three mutants in series, though the process might be quicker* (Haldane 1957). Therefore, the results presented in the current study should not be seen as a

bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

refutation of Haldane's ideas. Instead, we simply ask if the arguments developed by Haldane for the accumulation of costs over several loci in a clonally-reproducing population also apply to non-clonal, i.e., sexual species. Our conclusion is that sexual outbreeding and recombination allows the population to avoid such an accumulation of selective costs.

Materials and Methods

We performed numerical simulations of populations containing 100,000 diploid individuals. We did ten replicates of each simulation and the results are shown in the Figures below. We assumed linkage equilibrium in the initial population. We compared the outcome of multilocus selection in asexual and sexual populations. For the sexual populations, loci are spaced one centimorgan apart on a linear chromosome. The initial frequency of the favored allele is 0.01 at each locus. The fitness advantage of a favored allele is 0.05 and fitness interactions are multiplicative between loci. Individuals produced an average of two offspring each and selection occurred through differential viability that was dependent on the genotypic fitness value. We did a series of simulations in which we varied the number of loci under selection. In each case, we compared the outcome of selection in asexual and sexual populations. Following Haldane's example, we assumed that the favorable alleles were already segregating at a low frequency in the initial population, and that the population was in linkage equilibrium.

Results

The results described below illustrate two main points. First, we show why it was necessary for Haldane (1957) to implicitly assume a larger and larger initial population as the number of loci under selection increased. The reason was that Haldane's model did not include recombination between the selected loci; therefore, the initial population had to be large enough to contain at least one individual that contained all of the favorable mutations in its genotype. Secondly, we show that there is no need to increase the initial

bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

population size for multilocus selection in a sexually outbreeding population. This is because recombination will automatically produce the genotype with the maximum number of favored alleles later on during the selection process (see Discussion below).

We first performed simulations of the situation where there is only one locus undergoing selection. This allowed us to check the accuracy of our simulation method against the theoretical predictions of classical population genetics and it also provided a reference point for the subsequent simulations involving more than one locus. The results are presented in Figure 1. The results of ten replicate simulations are shown in Figure 1A, and the expected rise in allele frequency based on classical population genetics theory (Li 1955 page 258) is shown in Figure 1B. From the Figure, we see that the simulation results match the theoretical prediction very well and that it takes approximately 200 generations for the favored allele to rise from its initial frequency of 0.01 to fixation.

Having established that the simulation provides the expected results for selection at a single locus, we then performed simulations of selection acting at multiple loci in both sexual and asexual populations. The results for two loci are shown in Figure 2, Panels A and B. We see that the asexual and sexual populations yield different results, even when we are dealing with only two loci. This can be explained as follows. Given that the initial frequencies of the favored alleles at the two loci are 0.01, then the expected frequency of chromosomes carrying both favored alleles is one in ten thousand. Thus, we expect to see twenty such chromosomes among the 100,000 diploid genotypes in the initial population. But the expected frequency of diploids carrying two such chromosomes is only one in a hundred million; therefore, we do not expect to see such double homozygotes in the initial population. This means that the most favored combination in the initial population is expected to be either a doubly mutant chromosome in combination with a singly mutant chromosome, or a doubly mutant chromosome in combination with a non-mutant chromosome, or two singly mutant chromosomes. This explains why selection on the asexual population results in diploids carrying only two – or at most three - out the possible four favorable alleles that would occur in a double homozygote (see Figure 2, Panel A). Of course, if we had assumed a higher initial frequency for the favored alleles – say 0.1 rather than 0.01 – then the initial population would have contained some double homozygotes of the favored combination.

bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

On the other hand, the problem would have become even more acute if we had used lower initial allele frequencies. In contrast to the asexual population, however, the sexual population can generate diploid genotypes carrying four favorable alleles as the allele frequencies rise at the individual loci (see Figure 2B). In other words, recombination compensates for the absence of the optimal genotype in the initial population.

The results for selection acting simultaneously at four loci are shown in Figure 2, Panels C and D. In this case, we begin to see an even greater difference between the asexual and sexual populations. The reason for this difference is that the initial asexual population lacks many high fitness genotypes. And the reason for this absence is that the expected frequency of such individuals – given the initial allele frequencies of 0.01 and linkage equilibrium – is exceedingly small. For example, the expected frequency of diploid genotypes which are homozygous for the favorable allele at all four loci is only 10^{-16}. Consequently, it would require a population size of many trillions of diploid individuals in order to see one individual of the optimal genotype within the initial population. From Figure 2c we can see that the fittest genotypes in the asexual population contained only two or three out of the eight possible favorable alleles – giving average frequencies after selection of 0.25 and 0.375, respectively. Of course, the sexual population is also lacking these high fitness genotypes initially, but this is not a problem because random recombination can assemble such genotypes as the individual allele frequencies rise in response to selection. For example, once the allele frequencies have reached a frequency of 0.5 in an outbreeding sexual population we expect to see some individuals with the optimal genotype begin to appear within the population of 100,000 diploid individuals.

Finally, the results for selection acting on 100 loci are shown in Figure 2, Panels E and F. In this case, there is a striking difference between the asexual and sexual populations. The initial population contains only low fitness genotypes where we expect an average of only one favorable allele per chromosome – given that there are 100 loci and a frequency of 0.01 for the favorable allele at each locus (see Hickey and Golding 2018). The asexual population is again limited to selecting among such low-fitness genotypes. In contrast to this, the sexual population can use recombination to gradually assemble chromosomes with higher and higher numbers of favorable alleles as the

bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

frequencies of the favorable alleles rise in response to selection. At the end of the selection period, the sexual population consists of individuals that are homozygous for the favored allele at all one hundred loci.

These results show that, in a sexually outbreeding population, the frequency of favorable alleles at many loci can respond simultaneously to independent selection at rates that are very similar to the predicted rate at a single locus (compare Figure 2f and Figure 1b). This implies that the rate of response to selection at any given locus is not greatly impeded by selection at other loci. In other words, many loci can respond simultaneously to independent selection. In summary, our results show that an asexual population cannot respond to multilocus selection as efficiently as a sexual population because it cannot generate higher-fitness genotypic combinations as the allelic frequencies increase in response to selection. Moreover, the relative disadvantage of the asexual population increases as the number of loci under selection increases.

Discussion

In contrast to Haldane's prediction (Haldane 1957, 1960) our simulation results indicate that allele substitutions at many loci can happen as quickly as substitutions at a single locus – but only in a sexual population (see Figure 2, Panels B, D and F). Moreover, this rapid evolution can occur even when the reproductive excess averages only two offspring per individual (as was the case in our simulations). How can one explain this seeming contradiction? Clearly, Haldane's scenario does not explain our simulation results as shown in Figure 2. The reason for the discrepancy is, we believe, that Haldane's argument assumed that the optimal genotypic combination had to occur in the initial population and that genotypes can be considered as fixed entities that reproduce themselves from one generation to another. But these assumptions do not hold for a sexually outbreeding population. First, the population is usually not large enough to contain even a single individual with the vanishingly rare optimal genotypic combination; this fact has already been pointed out by Ewens (1972). Secondly, specific genotypes are broken down by recombination each generation and replaced by other genotypic combinations. The process of recombination involves the continual disassembly of

bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

existing genotypes and reassembly of new genotypes from one generation to the next. Because of this process, however, those genotypes which are expected to be vanishingly rare when the allele frequencies are low will be automatically generated through recombination once the allele frequencies begin to rise in response to selection (Muller 1932, Hickey and Golding 2018). Consequently, there is no need for the massive culling that would be required in a very large initial population that contained an extremely rare optimal genotype. Thus, in an outbreeding sexual population, recombination solves the perceived problem of costs that are cumulative over different genetic loci. In Haldane's scenario, it is necessary to "grow" the optimal genotypic combination from an initial, vanishingly rare frequency to fixation. In practice, however, this optimal combination is produced by recombination only near the end of the process. This avoids the huge cumulative cost.

Haldane's theory is mathematically very elegant; what happens in nature is mathematically more muddled, but it is much more efficient. In nature, selection favors a wide variety of suboptimal combinations which are then recombined. In this way, recombination can harvest all favorable mutations from any genetic background. It is often not until quite late in the selection process that the optimal genotypic combination first appears. For example, even after the allele frequencies at each of 100 loci have risen to a frequency of 0.9, the probability of randomly generating a chromosome with a favorable allele at all 100 loci is still less than 0.0001. By the time the individual alleles have reached a frequency of 0.99, however, more than one third of all chromosomes in the population are expected to contain the optimal combination - i.e., a favorable allele at all 100 loci. In other words, rather than increasing slowly from an infinitesimally low initial frequency as suggested by Haldane (1957), the optimal combination is generated through recombination near the end of the selection process.

The strategy used by nature is somewhat analogous to the process of parallel computing. Since recombination decouples the allelic changes at a given locus from changes at other loci, it effectively decomposes multilocus selection into several simultaneous instances of individual selection. Haldane (1957) envisaged that natural selection would have to do an exhaustive search through all possible genotypic combinations; instead, what happens in nature is a massively parallel heuristic search.

bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The process of alternating selection and recombination enables the sexual population to trace a relatively narrow path through the myriad of possible genotypic combinations until, eventually, it arrives at the optimum genotypic combination. A comparison between natural selection and parallel computing has also been made by Wilf and Ewens (2010) although those authors were referring to the sequential fixation of favorable alleles during long-term evolution, rather than to the effects of recombination within a population.

In a genetically polymorphic sexual population multilocus genotypes are highly ephemeral. The alternate rounds of genotypic sifting (by selection) and shuffling (by recombination) result in a gradual increase in the number of favorable alleles per chromosome. Eventually, chromosomes that contain a favorable allele at every locus appear. Thus, the genotype is the unit of selection in any given generation, but it is not the unit of evolution over many generations. Recombination dissolves individual genotypes each generation and recombines them into new genotypic combinations. In this way, recombination translates genotypic selection into genic evolution. This effect of recombination is not so obvious when there are only a few possible genotypic combinations. But when we consider many loci simultaneously, then each genotype is effectively unique. This is because existing genotypes are merely a small random sample of a much larger set of possible genotypes (Ewens 1972, Edhan 2017). Human populations provide a familiar example: every human individual is genetically distinct, not only from the billions of other individuals within the current population, but also from individuals in past and future populations. The genetic uniqueness of human individuals reflects the fact that the number of possible human genotypes is many orders of magnitude lager than the total number of humans alive today.

If the population is in linkage equilibrium, the expected number of favorable alleles per chromosome can be inferred from the current allele frequencies at the individual loci. But if all chromosomes contained exactly the expected number of favorable alleles, natural selection would be much less efficient. Selection acts on the distribution of favorable allele numbers around the expected mean. In this way, the coin-flipping aspect of particulate inheritance in a Mendelian population acts synergistically with natural selection. And this distribution of allele numbers per chromosome is regenerated every generation through recombination. This provides another perspective

bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

on the synergy between recombination and selection. This point is illustrated in Figure 3. Since both the asexual and sexual populations were initiated in linkage equilibrium, they had the same initial range of genotypic fitnesses. Because of this, the response to selection was similar in both populations in the early generations. As selection proceeded, however, the genotypic variation in the asexual population was depleted and the average fitness was limited to the highest fitness in in the initial population (see Figure 3a). In contrast to this, the sexual population can replenish its genotypic variation every generation through recombination. This causes the fitness range to be continuously shifted upwards, thus allowing selection to proceed far beyond the fitness range in the initial population. What is perhaps surprising is the magnitude of the selective advance in the sexual population compared to the limited response to selection in the asexual population, especially when we bear in mind that Figure 3 shows the results for only the first 60 generations of the simulation. The continued rise in the average frequency of the favored alleles over 200 generations (shown in Figure 2f) illustrates how the fitness of the sexual population continues to increase over 200 generations of selection.

When considering multi-locus selection, the huge potential fitness difference between the fittest possible genotype and the least fit genotype is often seen as unrealistically large. For example, given a selective advantage of only 0.05 per favoured allele and multiplicative fitness interactions between loci, a diploid genotype that is homozygous for the favoured allele at one hundred loci would have a fitness value that is more that 200-fold larger than the fitness of a genotype that is homozygous for the alternative allele at all 100 loci. The generation of such a huge fitness range in a sexual population has been seen as a problem (see Crow 1970). But this huge fitness range never occurs in an actual sexual population of, say, 100,000 individuals. In such a population, when the allele frequencies are low, the majority of the genotypic fitnesses also cluster around a relatively low value. And as the allele frequencies rise to higher values in response to selection, the fitnesses of genotypes also cluster around a correspondingly higher value. Because of recombination, the sexual population can traverse the very wide range of possible fitnesses without ever having an unrealistically large variance in fitness within the population at any given time. As the allele frequencies rise in response to

bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

selection, the genotypic fitness distribution moves as a relatively narrow travelling wave that traverses

bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the entire range of possible fitnesses. The early phase of this process is shown in Figure 3. Although there is a potential two-hundred-fold range in fitness between the theoretically least fit genotype and the theoretically most fit genotype, the range of fitnesses in the actual population at any given time is much less than this.

Maynard Smith (1971) asked "what use is sex?" Our answer to that question is that recombination avoids the potentially increasing costs of natural selection as the number of genes under selection increases. This answer was already hinted at in the literature, ever since Fisher (1930) stated that the advantage of sexual reproduction is proportional to the number of genetic loci under selection. Muller (1932, 1964) developed Fisher's idea and concluded that recombination among many loci could allow a sexual population to evolve "hundreds to millions of times faster" than a comparable asexual population. This same idea of an increasing advantage for recombination as the number of loci under selection increases, has been echoed in later papers (e.g., Crow and Kimura 1965, Otto and Barton 2001, Iles et al 2003, Park and Krug 2013, Edhan et al 2017). Park and Krug (2013) have shown that the advantage of recombination is proportional to L, where L is the number of loci under selection. Thus, the increasing advantage of sex as the number of loci under selection increases offsets the increasing cost of natural selection that was predicted by Haldane (1957).

It is interesting to note that Haldane (1957) did not see the relevance of Muller's ideas about recombination to the problem of reproductive costs that accumulated over loci. It is perhaps even more interesting that Muller (1964) also did not see the connection between his own ideas and the problem posed by Haldane (1957). But neither did their colleagues at the time see the connection (e.g. VanValen 1963). Indeed, we believe that we are the first authors to have pointed out the direct link between the ideas of Muller (1932) and those of Haldane (1957). And we have to admit that it was only because of the counterintuitive nature of our own simulation results that we were forced to re-examine Haldane's argument in the light of recombination. One could say that the solution to Haldane's Dilemma has been "hiding in plain sight" for six decades.

Our results support models of adaptation based on the standing genetic variation (see Barrett and Schluter 2007, Orr 2005, Reznick 2016, Schrider and Kern 2017). As noted by Messer et al (2016) rapid phenotypic evolution suggests that many genomic loci

bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

may contribute to strongly selected traits. Molecular studies have confirmed that indeed several loci are generally involved during rapid evolution (Burke et 2010, Teotonio et al 2009). Moreover, there is direct experimental evidence that recombination facilitates the rapid response to multilocus selection (McDonald et al 2016, Kosheleva and Desai 2018). The relevance of our work is that it removes any concerns about the high cumulative costs of multilocus selection as argued by Haldane (1957), provided of course that there is recombination occurring between the selected genes.

In this study, we assumed – for computational convenience – that there were only 100 biallelic loci. The early, conservative estimates of genetic variation in natural populations concluded that approximately twenty percent of the loci within the eukaryotic genome are polymorphic (Lewontin and Hubby, 1966). If we assume that there are 4,000 polymorphic loci – with two alleles at each locus – there are $2^{4,000}$ possible genotypic combinations, a number which is greater than $10^{1,000}$. This number is far greater than the total number of atoms in the observable universe – which has been estimated to be approximately 10^{80} (see Hogan 2000). Darwin stated that “endless forms” have evolved as a result of natural selection. We could say that this process has been facilitated through the production of effectively endless genotypic combinations as a result of sexual reproduction and recombination.

In conclusion, sex allows many genes to respond to independent selection at the same time. In a sexual population, selection on one gene does not slow the rate of adaptive response at another genetic locus. This is because recombination allows allelic change at one locus to become independent of the changes at other loci. As stated by Muller (1932), the allele frequencies at the different loci “diffuse through one another”. Consequently, even those organisms with limited population sizes and modest reproductive excess can evolve efficiently. Haldane (1957) predicted that the cost of selection would pose a particularly difficult challenge for such species. Perhaps this explains why sex tends to be most common among species with limited population sizes and modest reproductive excess.

Sexual reproduction could be described as “a solution in search of a problem” whereas Haldane's Dilemma is a problem in search of a solution. In this study, we show that the former can provide a solution for the latter.

bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Acknowledgements

This work was supported by a Discovery Grant from NSERC Canada to GBG (grant no. 140221-10). We are grateful to Richard Lewontin for stimulating discussions during the early development of this work. We would like to thank Chi Kuen Cheung for his support and encouragement.

References

Barrett, R.D.H., Schluter, D. 2007. Adaptation from standing genetic variation. Trends Ecol. Evol. 23: 38-44.

Burke, M.K., Dunham, J.P., Shahrestani, P., Thornton, K.R., Rose, M.R., Long, A.D. 2010. Genome-wide analysis of a long-term evolution experiment in Drosophila. Nature 467: 587-590.

Crow, J.F. 1970. Genetic loads and the cost of natural selection. In: Kojima, K-I. (Ed.), Mathematical Topics in Population Genetics. Springer, 128-177.

Crow, J.F., Kimura, M. 1965. Evolution in sexual and asexual populations. Amer. Natur. 99: 439-450.

Dodson, E.O. 1962. Note on the cost of natural selection. Amer. Natur. 96: 123-126.

Edhan, O., Hellman, Z., Sherill-Rofe, D., 2017. Sex with no regrets: How sexual Reproduction uses a no regret learning algorithm for evolutionary advantage. J Theor Biol. 426, 67-81.

Ewens, W. J. 1972. The substitutional load in a finite population. Amer. Natur. 106: 273-282.

Felsenstein, J. 1971. On the biological significance of the cost of gene substitution. Amer. Natur. 105: 1-12.

Fisher, R.A., 1930. *The Genetical Theory of Natural Selection*. Oxford: Clarendon.

Grant, V., Flake, R.H. 1974. Population structure in relation to the cost of selection.

bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Proc. Nat. Acad. Sci USA 71: 1670 -1671.

Haldane, J.B.S. 1957. The cost of natural selection. J. Genet. 55: 511-524.

Haldane, J.B.S., 1960. More precise expressions for the cost of natural selection. J. Genet. 57: 351-360.

Hickey, D.A., Golding, G.B. 2018. The advantage of recombination when selection is acting at many loci. J. Theoret. Biol. 442: 123-128.

Hogan, C.J. 2000. Why the universe is just so. Rev. Mod. Phys. 72: 1149-1161.

Iles, M.M., Walters, K., Cannings, C., 2003. Recombination can evolve in large finite populations given selection on sufficient loci. Genetics 165, 2249-2258.

Kern, A.D. and Hahn, M.W. 2018. The neutral theory in light of natural selection. Mol. Biol. Evol. 35: 166-1371.

Kimura, M. 1968. Evolutionary rate at the molecular level. Nature 217: 624-626.

Kimura, M. 1995. Limitations of Darwinian selection in a finite population. Proc. Nat. Acad. Sci. USA 92: 2343-2344.

Kosheleva, K., Desai, M.M., 2018. Recombination alters the dynamics of adaptation on standing variation in laboratory yeast populations. Mol. Biol. Evol. 35: 180-201.

Lewontin, R.C., Hubby, J.L., 1966. A molecular approach to the study of genic Heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of Drosophila pseudoobscura . Genetics 54, 595-609.

Li, C.C., 1955. *Population Genetics*, p 258. (University of Chicago Press).

Maynard Smith, J. 1968. “Haldane’s dilemma” and the rate of evolution. Nature 219: 1114-1116.

Maynard Smith, J. 1971. What use is sex? J. Theoret. Biol. 30: 319-335.

McDonald, M. J., Rice, D. P., Desai, M. M., 2016. Sex speeds adaptation by altering the dynamics of molecular evolution. Nature 531, 233-236.

Messer, P.W., Ellner, S.P., Hairston, N.G. 2016. Can population genetics adapt to rapid evolution? Trends Genet. 32: 408-418.

Muller, H.J., 1932. Some genetic aspects of sex. Amer. Natur. 66: 118-138.

Muller, H.J., 1964. The relation of recombination to mutational advance. Mutat. Res. 1: 2 - 9.

bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Orr, H.A. 2005. The genetic theory of adaptation: a brief history. Nat. Rev. Genet. 6: 119-127.

Otto, S.P., Barton, N., 2001. Selection for recombination in small populations. Evolution 55, 1921-1931.

Park, S.-C., Krug, J. 2013. Rate of adaptation in sexuals and asexuals: a solvable model of the Fisher-Muller effect. Genetics 195: 941-955.

Reznick, D. Hard and soft selection revisited: how evolution by natural selection works in the real world. J. Hered. 107: 3-14..

Schrider, D. R. and Kern A.D. 2017. Soft sweeps are the dominant mode of adaptation in the human genome. Mol. Biol. Evol. 34: 1863-1877.

Sved, J.A. 1968. Possible rates of gene substitution in evolution. Amer. Natur. 102: 283-293.

Teotónio, H., Chelo, I.M., Bradic, M., Rose, M.R., Long, A.D., 2009. Experimental evolution reveals natural selection on standing genetic variation. Nat Genet. 41, 251-257.

VanValen, L. 1963. Haldane's dilemma, evolutionary rates, and heterosis. Amer. Natur. 97: 185-190.

Wilf, H.S., Ewens, W.J., 2010. There's plenty of time for evolution. Proc. Nat. Acad. Sci. USA 107: 22454-22456.

bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

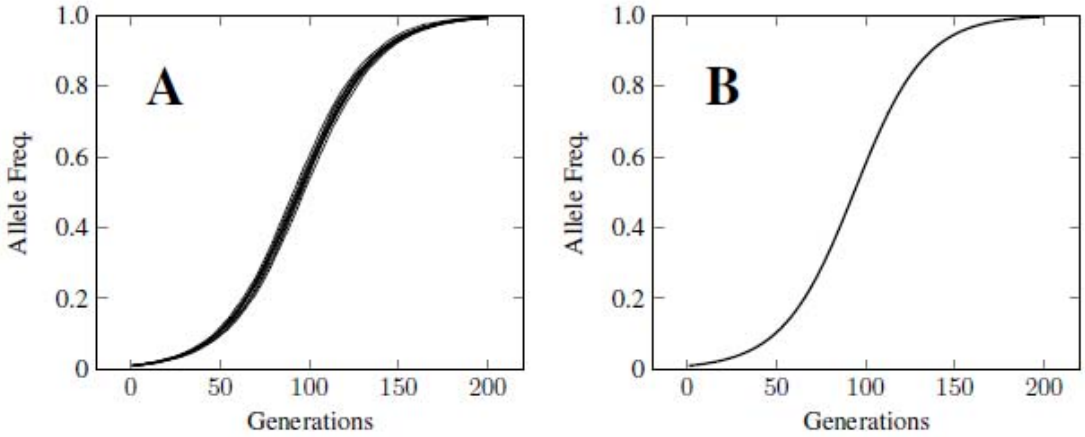
Figure 1. Selection at a single locus, given an initial frequency of 0.01 and a selective advantage of 0.05. The population size was 100,000 diploid individuals. Panel A Shows simulation results for 200 generations (10 replicates, results for individual replicates are shown). Panel B shows the theoretical, predicted rise in frequency of a favored allele at a single locus given an initial frequency of 0.01 and a selective advantage of 0.05.

Figure 2. Selection at multiple loci, given an initial frequency of 0.01 for the favorable alleles and a selective advantage of 0.05. The graph shows the rise in the average frequency of favorable alleles during the course of selection. The reason that this frequency does not reach a value of 1.0 in the asexual populations is because the highest fitness genotype in the initial populations did not contain the maximum number of favorable alleles (see text). Panels A and B show the results for asexual and sexual populations, respectively, when selection is acting at two loci. Results for each of ten replicate simulations are shown separately. Panels C and D show comparable results for four loci. Panels E and F show the results for 100 loci.

Figure 3. The rise in fitness in response to selection in sexual and asexual populations. Panel A shows the results for an asexual population of 100,000 asexual individuals; Panel B shows the results for an outbreeding sexual population of the same size. The solid line shows the average fitness and the dashed lines show the standard deviation in fitness. The horizontal dotted line in Panel A indicates the highest fitness in the initial asexual population. There were 100 loci subject to selection; the initial frequency of favorable alleles was 0.01 at each locus; the selective advantage of a favorable allele was 0.05.

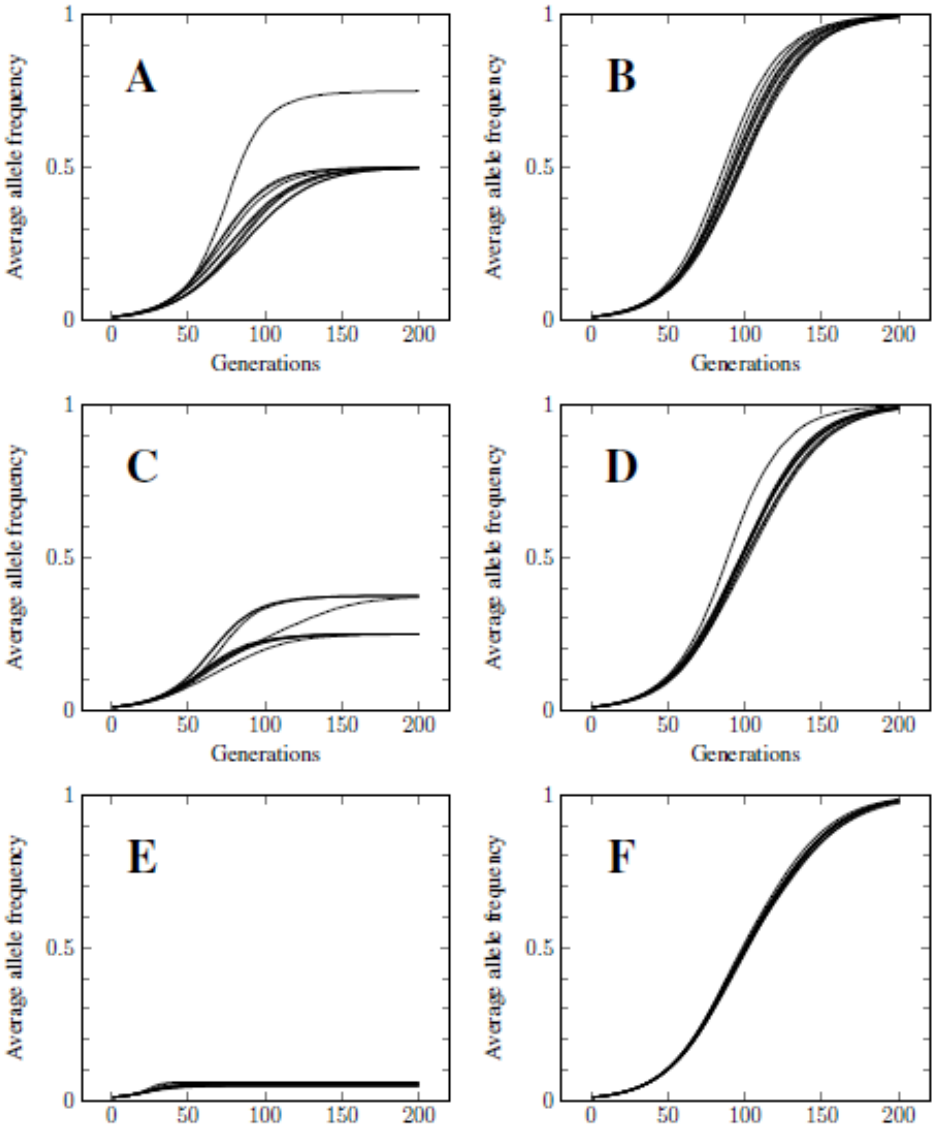
bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

FIGURE 1



bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

FIGURE 2



bioRxiv preprint doi: https://doi.org/10.1101/582536; this version posted March 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

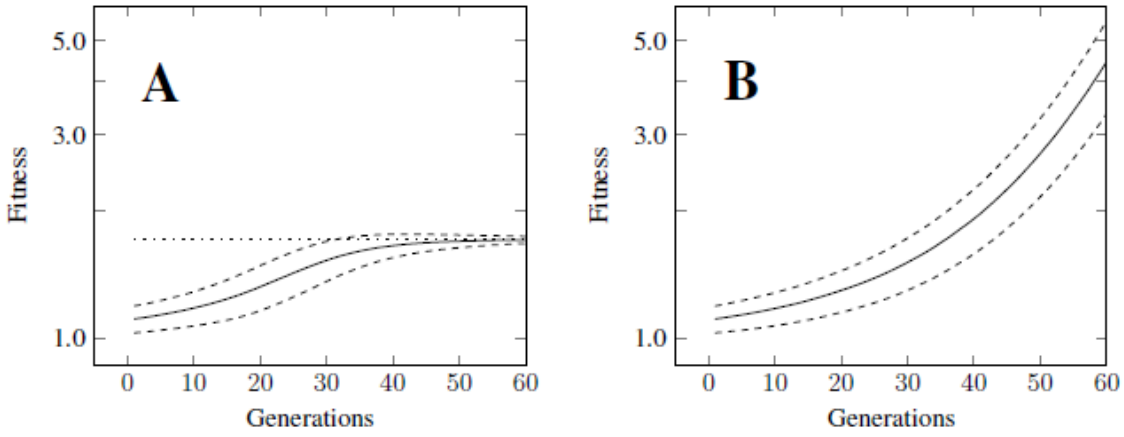



FIGURE 3