Finite Populations, Finite Resources, and the Evolutionary Maintenance of Genetic Recombination

SOL ACKERMAN, AMIR R. KERMANY, AND DONAL A. HICKEY

From the Department of Biology, Concordia University, 7141 Sherbrooke West, Montréal, Québec, Canada, H3G 1M8.

Address correspondence to Donal A. Hickey at the address above, or e-mail: dhickey@alcor.concordia.ca.

Abstract

Most previous models for the evolution of sex implicitly assume infinite population sizes and limitless resources. However, because favorable mutations are very rare and eukaryotic populations are finite, it has already been shown that multiple favorable mutants virtually never occur by chance. Therefore, sex is required to combine different favorable mutations into a single lineage. Second, we show that even when multiple favorable mutations do coexist, competition between genotypes can create negative epistasis for fitness, thus favoring recombination. Competition is especially effective when selection is at the level of viability in K-selected species living in a resource-limited environment. This means that recombination is advantageous both for incorporating new favorable mutations into the gene pool and for accelerating their increase to fixation. These advantages of recombination are diminished, however, as genome sizes decrease or as the amount of competition within the species is a less important component of selection.

Key words: competition, epistasis, evolution of sex, fitness, recombination, sex

Sex is not a necessity of life; this is clearly illustrated by the many instances of asexual reproduction that exist, especially among microbes. Yet, most biologists believe that sexual reproduction and the resulting genetic recombination do provide some selective advantage and, based on the broad patterns of distribution of sex among taxa, we can infer that the advantage of recombination is more important for large multicellular eukaryotes than it is for small single-celled prokaryotes: recombination and sexual reproduction are ubiquitous among higher eukaryotes (Hadany and Comeron 2008), whereas bacteria are primarily asexual (Narra and Ochman 2006). But we still do not know what the nature of this advantage is. Indeed, the evolution of sex is one of the longest standing unsolved problems in biology (for a review, see Agrawal 2006). It is now almost 80 years since Ronald Fisher wrote that sexual reproduction “is a development of some special value to the organisms which employ it” (Fisher 1930). Eight decades later, we are still trying to decipher what that “special value” might be. As Sarah Otto remarked recently (see Otto 2009), we are still trying to “solve the paradox of sex.”

Biologists have invested a large amount of research effort into this problem and several models for the evolution of genetic recombination have been proposed. Earlier theories proposed that the benefit of recombination lay in the fact that it combined different beneficial mutations into a single genotype (Muller 1932, 1964; Crow and Kimura, 1965, 1970). Later work, however, showed that the benefits of combining advantageous mutations are offset by the fact that recombination also breaks up these favorable combinations once it has been formed (Maynard Smith 1968, 1978). In other words, recombination simply randomizes genotypes, without regard to the fitness of the alleles being recombined.

Population genetics theory predicts that genetic recombination will be an advantage in those situations where there is a dearth of favorable combinations in the population, that is, there is negative linkage disequilibrium between alleles at selected loci (Maynard Smith 1978). In this case, simple randomization of genotypes will increase the variance in fitness among the genotypes and thus accelerates the response to selection. Negative linkage disequilibrium could arise either through the effects of genetic drift (Hill and Robertson 1966; Felsenstein 1974; Keightley and Otto, 2006) or through the action of selection itself, provided that selection acted in a negatively epistatic fashion (Kimura and Maruyama 1966; Kondrashov 1982, 1988; Barton 1995). The challenge for evolutionary biologists and ecologists, however, has been to describe conditions that are both widely distributed in nature and that would result in negative epistasis for fitness. Spatially complex environments, such as described by the Tangled Bank model (Bell 1982), might be expected to favor sexual organisms that produce genetically variable offspring, but it is
temporally fluctuating environments—such as described by the Red Queen model (Van Valen 1973; Hamilton 1980; Hamilton et al. 1990; Peters and Lively 1999)—that provide the conditions necessary for the buildup of negative linkage disequilibrium between loci. Not only does selection have to fluctuate over time, however, it has to undergo periodic reversals. This has been most extensively studied in the case of host–parasite coadaptation cycles. Although the Red Queen explanation for the evolution of sex is attractive, its general relevance to the problem of sex is still debated (Neiman and Koskella 2009) because it depends on particular ecological conditions, such as host–parasite coevolution. In this situation, the adaptation of parasites to the most common host genotypes will provide an advantage to hosts with a rare genotype produced by recombination. This advantage, by its very nature, is negatively frequency dependent. Consequently, as the frequency of favored rare genotype increases, the direction of selection will be reversed. However, it has been shown that the periodicity of the selection reversals is quite constrained (Otto and Nuismer 2004; Salathé et al. 2008). Thus, although this form of selection could explain the evolution of sex, some biologists feel that it requires a set of ecological circumstances that seem too narrow to explain the near-universality of sexual reproduction among eukaryotes.

Our approach has been to concentrate on the effects of competitive interactions between genotypes within a population. Whereas there is a large theoretical literature describing the interactions between different species in an ecosystem, relatively little attention has been paid to competitive interactions between genotypes within a population. Instead, genotypes are usually assigned fixed fitness values that do not depend either on the availability of resources or on the frequencies of other genotypes in the population. Yet, as noted by Lewontin (1955), “it would be strange if what applied to different species did not apply to some extent to different genotypes within the same species.” Thus we have asked how selection would play out if the genes affected fitness indirectly through the production of fixed phenotypes with different competitive abilities. In this scenario, the phenotype is uniquely determined by the genotype, but the resulting fitness depends not only on the phenotype itself but also on the phenotypes of the competing genotypes. This scenario should apply to any species where the “struggle for existence” not only involves a struggle with the external environment but also some degree of competitive struggle with other genotypes within the same species. Most previous models of intraspecific competition between genotypes have focused primarily on their potential for maintaining genetic polymorphism (Trotter and Spencer 2007) or for promoting speciation (Bürger et al. 2006). A couple of other studies, however, have looked explicitly at the relationship between intraspecific competition and recombination (e.g., Peck and Waxman 2000; Bagnoli and Guardiani 2005).

**Finite Populations**

Population genetics models tend to fall into 2 categories: deterministic models that imply infinite population size and stochastic models that emphasize the effects of gamete sampling in small populations. But many natural populations are neither very small nor infinitely large; they exist in a gray area where we are not sure whether their evolution is dominated by deterministic or stochastic forces. For many years, the field of population genetics was engulfed in a controversy between those who believed that stochastic sampling effects were insignificant and those who believed that, to the contrary, the random sampling of gametes in finite populations was a major force behind the change in allelic frequencies. The deadlock has been broken by the accumulation of a large amount of comparative DNA sequence data. From those data it is clear that stochastic sampling is a major force in DNA sequence divergence. This means that the population size, \( N \), is generally smaller than the reciprocal of the mutation rate, \( \mu \). Roughly speaking, we can conclude that real population sizes, although they may be large, are generally significantly less than one billion individuals. This finding is important for discussions of the evolution of genetic recombination, especially when we consider the expected numbers of double mutations in natural populations. For example, Maynard Smith (1968, 1978) pointed out that double mutants could be produced by mutation alone provided that the population size were large enough, making the generation of doubly favorable mutants by recombination unnecessary. But for this to happen in practice, the population size would have to be of the order of \( 10^{16} \), and for triple mutants we would need population sizes of \( 10^{24} \). These numbers are based on the estimate of \( 10^{-3} \) favorable mutations per genome per generation, as reported by Perfeito et al. (2007), which translates into a per-gene rate of \( 10^{-8} \), assuming that there are at least 1000 genes per haploid genome. The projected population numbers are clearly unrealistic for the majority of eukaryotic species, especially multicellular eukaryotes. Although it is often impossible to measure population sizes directly, it is also impossible to measure historical effective population sizes, we can obtain indirect measures because the effective size \( N_e \) of a population is related to the genetic diversity that is maintained in the population (Kimura and Crow 1964). Such analyses of polymorphism data yield an estimated effective population size of approximately 10 000 for the modern human lineage (e.g., Takahata 1993; Yu et al. 2001) and less than 100 000 for other primate lineages (Burgess and Yang 2008). Given that the mutation rate for favorable alleles is orders of magnitude less than the reciprocal of the population size, new favorable mutants in a given gene are essentially unique and are, for all practical purposes, nonrecurring. It should be noted that this problem is less acute for prokaryotes and single-celled eukaryotes that can have very large population sizes and that have rapid doubling times (20 min in *Escherichia coli* vs. 20 years in humans). Consequently, without recombination most of the favorable mutants that occur in finite populations—even very large finite populations—will eventually be lost. For instance, Barrick and Lenski (2009) showed that, in asexual populations, some beneficial mutations can be lost through competition with other beneficial mutations.
**Finite Resources**

In the previous section, we have tried to demonstrate why the larger genomes, smaller population sizes, and slower replication rates that characterize multicellular eukaryotes, as compared with many single-celled prokaryotes, might favor the maintenance of genetic recombination. A correlated difference between large multicellular eukaryotes and small single-celled prokaryotes is the greater importance of competition among large eukaryotes. In ecological terms, prokaryotes tend to be r-selected, whereas multicellular eukaryotes are more K-selected—although this is not a universal rule of course. Although it is difficult to measure competition in absolute terms, a number of predictions have been made regarding the response of natural selection to increasing levels of intraspecific competition. For instance, Adler and Levins (1994) predicted that viability selection would become more important than fecundity selection as the level of intraspecific competition increased, and a number of subsequent studies (e.g., Mappes et al. 2008) have provided empirical evidence in support of this prediction. Here, we explore how the competitive component of selection also favors the maintenance of genetic recombination.

Natural selection favors genotypes that are better adapted to the environment; this includes both the physical and biotic environments. In the case of the biotic component of the environment, most researchers have focused on interactions between species in an ecosystem and relatively little attention has been paid to the interactions between individuals within a species. Indeed, it is this competitive interaction between individuals that Darwin (1859) saw as the essence of his “struggle for existence.” He wrote: “Two canine animals in a time of dearth, may be truly said to struggle with each other which shall get food and live. But a plant on the edge of a desert is said to struggle for life against the drought, though more properly it should be said to be dependent on the moisture. A plant which annually produces a thousand seeds, of which on an average only one comes to maturity, may be more truly said to struggle with the plants of the same and other kinds which already clothe the ground.” In other words, Darwin saw natural selection as a zero-sum game where the number of surviving offspring is no greater than the number of parental individuals and where competition between conspecific individuals plays a crucial role. This is in contrast to many population genetics models that assign fixed fitnesses to genotypes which then implicitly engage in replication races during which less fit types are diluted out by fitter types rather than being culled from the population.

In our model, which mirrors Darwin’s scenario where only one offspring on average comes to maturity, the genotype specifies a fixed phenotype and this phenotype determines competitive ability. Fitness depends on the outcome of pairwise competition between individuals. This means that fitness depends not only on the phenotype of the individual itself but also on the phenotype of its competitor, and the latter depends in turn on the frequency of competitor genotypes. This is summarized in Table 1. Moreover, as selection proceeds, the frequency of the competitors changes; specifically, the frequency of the better competitors increases making competition more stringent. Meanwhile, since the population size does not increase, there is always an average of one surviving offspring per parent, yielding a constant average fitness, with a value of 1, for the population.

Based on the competitive interactions shown in Table 1, the expected fitnesses of the individual genotypes were calculated as follows:

\[ W_{AB} = \frac{P_{AB} + (P_{AB})(P_{Ab}) + (P_{AB})(P_{aB}) + (P_{AB})(P_{ab})}{P_{AB}} \]

\[ W_{Ab} = \frac{P_{Ab} - (P_{Ab})(P_{AB}) + (P_{Ab})(P_{aB})}{P_{AB}} \]

\[ W_{ab} = \frac{P_{ab} - (P_{ab})(P_{AB}) + (P_{ab})(P_{aB})}{P_{AB}} \]

\[ W_{aB} = \frac{P_{aB} - (P_{aB})(P_{AB}) + (P_{aB})(P_{aB})}{P_{AB}} \]

where \((P_{AB})\) represents the frequency of the AB genotype, etc. among the competing offspring. (These equations are also shown in legend of Figure 1b).

Figure 1 illustrates how competitive selection compares to more traditional population genetics models. We have used an example of a haploid model with 2 biallelic loci. In the case of models with fixed genotypic fitnesses (Panel A), we see that the mean fitness of the population increases as the frequency of the fitter genotypes increases. In the competitive selection model, however (Panel B), where the average population fitness remains fixed with a value of one, the realized fitnesses of the individual genotypes change in a negative frequency-dependent fashion. Given that the individual fitnesses change in a frequency-dependent manner in the competitive selection model, we asked if the value of realized epistasis—which is calculated based on those fitnesses—also changed in a frequency-dependent manner. The answer to that question is answered in Panel C of Figure 1. From the Figure, we see that epistatic values are also frequency dependent in this model and, more important, they are negative. This shows that competitive selection between individuals within a population can generate negative linkage disequilibrium between the selected alleles, which would in turn favors genetic

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Genotype of competitor</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>Ab</td>
</tr>
<tr>
<td>0</td>
<td>+</td>
</tr>
<tr>
<td>Ab</td>
<td>–</td>
</tr>
<tr>
<td>aB</td>
<td>–</td>
</tr>
<tr>
<td>ab</td>
<td>–</td>
</tr>
</tbody>
</table>

For each of the 4 genotypes listed in the rows on the left, their success in competition depends both on their own phenotype and on the phenotype of the competing individual, listed in the columns at the top. Competitive success is indicated by a plus sign; failure by a minus sign; and a draw between equal competitors (50% chance of success or failure) is indicated by a zero.
recombination between the loci. We can illustrate this with a simple numerical example. If we assume that the frequencies of the favored alleles at each of the 2 loci are initially equal to 0.1, and that the genotypes are in linkage equilibrium, then the formulas in Figure 1b give values of 1.99, 1.80, 1.80, and 0.81 for the realized fitnesses of genotypes AB, Ab, aB, and ab, respectively. If we express these values as relative fitness, they become 1, 0.9, 0.9, and 0.41. Now if we use these relative fitnesses, in turn, to calculate the value of epistasis using the formula in Figure 1c, we get $E = -0.4$, that is, negative epistasis.

The calculations above provide us with the expected outcome of this type of competitive selection. In order to verify these expectations, we explored the effects of competitive selection by performing individual-based computer simulations. These simulations are not dependent on the formulas but on simple rules of competition. Nevertheless, the results of the simulations matched the expectations based on the equations.

Figure 1. Comparison between selection models with constant genotypic fitness (Panel A) and constant population fitness (Panel B).

Panel (A) The correlation between fitness and allele frequency when genotypes have constant fitness values.

The average fitness of the population is calculated as:

$$W_{bar} = (W_{AB})(P_{AB}) + (W_{Ab})(P_{Ab}) + (W_{aB})(P_{aB}) + (W_{ab})(P_{ab})$$

Where $W_{AB}$ represents the fitness of the AB genotype and $P_{AB}$ represents its frequency, etc.

Note that in this case, the average fitness of the population, $W_{bar}$, increases as the frequency of the favored allele increases. The line for $W_{aB}$ is not shown because it is identical to that shown for $W_{Ab}$.

Panel (B) The correlation between fitness and allele frequency when the total number of surviving offspring is constrained to equal the total number of parents.

The expected fitnesses of the individual genotypes were calculated as a function of their own frequency and that of other genotypes in the population of competing offspring, as follows:

$$W_{AB} = [P_{AB} + (P_{AB})(P_{Ab}) + (P_{AB})(P_{aB}) + (P_{AB})(P_{ab})] / P_{AB}$$

$$W_{Ab} = [P_{Ab} - (P_{Ab})(P_{AB}) + (P_{Ab})(P_{aB})] / P_{Ab}$$

$$W_{aB} = [P_{aB} - (P_{aB})(P_{AB}) + (P_{aB})(P_{ab})] / P_{aB}$$

$$W_{ab} = [P_{ab} - (P_{ab})(P_{AB}) - (P_{ab})(P_{aB}) + (P_{ab})(P_{Ab})] / P_{ab}$$

where $(P_{AB})$ represents the frequency of the AB genotype, etc. among the competing offspring. See text for an explanation of the formulas.

Panel (C) The correlation between epistasis and allele frequency, based on the realized fitness values shown in Figure 1b.

The value of epistasis was calculated as follows:

$$E = W_{AB}W_{ab} - W_{Ab}W_{aB}$$

Note that the fitness values that go into the calculation of epistasis are themselves changing in a frequency-dependent manner as shown in Figure 1b.
In the simulations individuals competed in pairs and the expected outcome of the pairwise competitive interactions was dependent on the phenotypes of the 2 competitors, as shown in Table 1. We assume a 2-locus, diallelic, haploid population of competing individuals with genotypes ab, Ab, aB, and AB. Generations are discrete and nonoverlapping, and the population has the carrying capacity $K$. At the beginning of each generation, each individual produces Poisson ($2$) number of offspring. The population is then reduced down to a size of $K$ through simultaneous pairwise competitions. Therefore, when the population has reached carrying capacity, we expect the population to produce $2K$ offspring, $K$ of which are subsequently eliminated through competition. When competing, the individual with the greatest number of uppercase alleles wins the competition (AB > Ab = aB > ab). We performed simulations with various population sizes and with different initial genotype frequencies. The results were consistent over a wide range of parameter values.

In the example shown in Figure 2, simulations were initiated with populations of 10 000 haploid individuals with 2 biallelic loci and where the frequencies of the favored alleles, A and B were initially low (0.1). Each favored allele increased the phenotypic value of the individual by 0.1, changing that value from 1 to 1.1. We assumed no epistatic interaction between loci. In other words, the phenotypic value of the AB genotype was assumed to be multiplicative and had a value of 1.21. Despite the lack of epistasis at the phenotypic level, however, epistatic effects at the fitness level were caused by the competitive selection itself. As pointed out by Milkman (1973), in a competitive selection situation, the phenotype represents the fitness potential rather than the fitness itself. The resulting realized fitness of each genotype is the result of an interaction between the fitness potential and the frequency of other competing genotypes in the population. It is the nonlinear mapping of fitness potential onto realized fitness that produces negative epistasis. This is why it is possible to have nonepistatic effects at the phenotypic level, but epistasis at the realized fitness level.

Genetic recombination (random assortment of alleles at the 2 loci) was included in 10 replicate simulations; the other 10 replicates lacked recombination. As we can see from Figure 2, those simulations that included a round of recombination resulted in a more rapid fixation of the doubly favored genotype, AB. The advantage of recombination is that it reduces the negative linkage disequilibrium that builds up in response to the negatively epistatic competitive selection (see Figure 1c).

The fact that fixation of the doubly mutant genotype, AB, occurs faster in the recombining population suggests that sexual strains would be at an advantage in a mixed population. We tested this prediction directly by simulating mixtures of sexual and asexual individuals within a single population. As shown in Figure 3, the frequency of the sexual type does indeed increase during the course of competitive selection. Our next step will be to test populations where all the individuals are sexual but where there is allelic variation for recombination rate.

**Discussion**

Our model shows that recombination can be advantageous provided that resources are limited. This assumption of limited resources results in a finite maximum population size
and it implies a significant degree of competition between individuals, especially once the population reaches its carrying capacity. Competition results in frequency-dependent realized fitnesses, and the epistatic values calculated from these fitnesses are themselves both frequency dependent and negative. This means that the competitive selection process itself creates negative linkage disequilibrium between selected alleles. The advantage of recombination is that it breaks down this negative linkage disequilibrium.

A number of previous studies have pointed out that sex may be advantageous in small populations (Felsenstein 1974; Christiansen et al. 1998; Otto and Barton 2001; Barton and Otto 2005) or in geographically subdivided large populations (Martin et al. 2006; Agrawal 2009). Our particular focus on population size relates particularly to the occurrence of new favorable mutations in multilocus genotypes. We argue that since actual population sizes are generally less than the reciprocal of the mutation rate per gene, and consequently orders of magnitude less than the favorable mutation rate, new favorable mutations are effectively unique events. Consequently, without recombination, the majority of these mutations will be lost from the population. This would not be the case if the population size was extremely large and the new favored mutant type had the capability of expanding to very large numbers of descendants in a short period of time. Such conditions may be met in bacterial populations but not in most multicellular animal populations. The problem for multicellular eukaryotes is compounded by the fact that they have genomes that are typically an order of magnitude greater in size than bacterial genomes. This combination of larger genomes, smaller population sizes, and slower replication rates represent special challenges for adaptive evolution in multicellular animals. Genetic recombination mitigates this problem to some extent.

The interaction between sex and competition was studied previously by Case and Taper (1986). In their model, which included both viability and fertility selection at the species level, the advantage of sex lay in the potentially broader niche of the sexual species, given an environmental gradient. Doncaster et al. (2000) also studied the evolution of sex in an ecological context and they concluded that intraspecific competition could lead to the coexistence of sexual and asexual species. Their model focused primarily on predator–prey systems. A single-species model was studied by Peck and Waxman (2000) focusing on the effects of scramble competition, with recurrent deleterious mutations affecting fertility only. Their model showed that recombination could be advantageous if individuals competed in small groups. Our model focuses on contest competition affecting viability (for a review of single-species competition models, see Brännstrom and Sumpter 2005). Our model does not require grouping of competitors; it simply assumes that total resources are finite—as is the total population size—resulting in the fact that for every winner there is a loser. In other words, in our model competition is a zero-sum game. The recent model proposed by Bagnoli and Guardiani (2005) is closer to what we present here in that it is a discrete, individual-based model, although there are also significant differences. For instance, their model is based on the framework of quantitative genetics, which means that they do not track the frequency of specific genotypes nor do they calculate values of epistasis. In addition, they consider an environmental gradient where competition is most intense between phenotypes adapted to the same point on that gradient. The fact that several previous studies, along with the current work, show an advantage for recombination in a competitive selection context, using a variety of different competition models, gives us some confidence that a wide range of competitive interactions may have this property. One attraction of our model is that, because of its simplicity, it makes the relationship between competition and negative epistasis more transparent.

Some other proposals for an advantage for sexual reproduction are implicitly based on some form of competition between conspecifics. These include theories of sib competition (Bulmer 1980) and clonal interference (Kim and Orr 2005). A recent study (Cooper 2007) provided direct experimental evidence that recombination can reduce clonal interference between selected favorable alleles in bacterial populations.

Conclusion

Many previous studies of the evolution of sex have implicitly assumed infinite populations that have access to limitless resources. Although such models may provide a good approximation for describing the evolutionary dynamics of experimental populations of bacteria growing in a chemostat, they are not realistic for many species, such as most vertebrates, where population sizes are limited and there is significant competition for limited resources. For such species, there could be a significant component of selection based on competitive ability with conspecifics. Considering competition also helps us to put the “cost of males” into perspective: in a competitive situation, producing 2 losers does not outweigh the benefits of producing one winner.

In summary, sex is not a necessity of life but, given limited population sizes and/or limited resources, it might give a competitive edge to those genotypes that recombine their genomes.

Funding

Discovery Grant from Natural Sciences and Engineering Research Council Canada (RGPIN 8516-2008) to D.A.H.; Canada Foundation for Innovation; Canada Research Chairs program.

Acknowledgments

This research was supported by a Graduate Scholarship (S.A.).

References


Corresponding Editor: Maurine Neiman

Received October 8, 2009; Revised January 15, 2010; Accepted January 26, 2010