A microscopic model of evolution of recombination

FRANCO BAGNOLI
Dipartimento di Energetica
(also INFN and CSDC),
Università di Firenze,
Via S. Marta 3, 50139 Firenze,
ITALY

CARLO GUARDIANI
Centro Interdipartimentale per lo Studio di
Dinamiche Complesse
Università di Firenze,
Via Sansone 1, 50019 Sesto Fiorentino,
ITALY

Abstract: - We study the evolution of recombination using a microscopic model developed within the frame of the theory of quantitative traits. Two components of fitness are considered: a static one that describes adaptation to environmental factors not related to the population itself, and a dynamic one that accounts for interactions between organisms *e.g.* competition. We focus on the dynamics of colonization of an empty niche. As competition is a function of the population, selection pressure rapidly changes in time. The simulations show that recombination provides a high velocity of movement in the phenotypic space thus allowing recombinants to colonize the highest fitness regions earlier than non recombinants that are often driven to extinction. The stabilizing effects of competition are also discussed.

Key-Words: - evolution, recombination, competition, assortativity, sex,environment

1 Introduction

One of the most challenging problems of evolutionary biology is that of the evolution of sexual reproduction, usually referred to as *The paradox of sex*. Sexual reproduction, in fact, is seemingly a very inefficient reproduction strategy, being associated to several costs, yet it is widespread in all taxa of the natural world. Actually [1], the only ancient, species-rich, fully asexual taxa are the Rotifer class *Bdelloidea* and the ostracod family *Darwinulidae*. The evolutionary success of sexual reproduction is paradoxical if we consider the several costs related to sex. First of all, there are costs related to attracting a partner, as the considerable resources invested by plants in floral display and nectar rewards. Moreover, the secondary sexual characters necessary to attract mates, as the bright colours of the males of many species of birds, make them more vulnerable to predators. Another disadvantage of sex is that it might break up favorable combinations of genes (that enabled the parents to survive to reproductive age) without any guaranty to produce better ones. A final, well-known problem is the *twofold price of sex*:in a population of N individuals, an asexual species will produce mN offsprings, while only $mN/2$ will be produced by a sexual species. Sexual reproduction therefore appears to be a very ineffective reproduction strategy leading to a short-term evolutionary disadvantage.

According to several recent studies syngamy originally evolved for trophic reasons [2, 3, 4, 5], but sex then fixed in the populations due to the several advantages of gene mixing. Recombination first of all increases the probability of fixation of beneficial mutations [6, 7, 8, 9] and decreases the probability of fixation of harmful mutations as in a recombinant

population the association of a mutation with the genetic background is only transient [10, 7, 11, 8]. A second theoretical advantage of recombination is that it brings together in the same genome several favorable mutations [12, 13, 14]. In non recombinant populations favorable mutations are segregated in different cell lines; only if two or more beneficial mutations occur independently in the same cell, it will be possible to find several favorable mutations in the same cell lineage but this event is very rare. Unfortunately, no experiments so far have been performed to test this hypothesis [15]. The third theoretical advantage of sex and recombination is that it increases genetic variability upon which selection can act. Theoretical arguments however, show that this is true only in the case of negative linkage disequilibrium and epistasis (the frequency and fitness of intermediate phenotypes is higher than that of the extreme ones) [16, 17]. Experimental studies however show that this situation is not very common so that another ingredient is apparently missing [15]. Recent studies seem to argue that the missing ingredient is a rapidly changing environment [18, 17]. A change in the environmental conditions in fact, can reverse the sign of epistasis *i.e.* the highest fitness group can shift from the intermediate to the extreme phenotypes and *vice versa*. In our work we tested the influence of a changing environment on the evolution of recombination. The change in environmental conditions is not imposed externally (changes of weather in the seasons of the year, or catastrophic changes like ice ages) but is related to the dynamics of the population itself through competition interactions. An increase in the number of competitors in fact, represents a deterioration in the environmental conditions that decreases the fitness of a given phenotype.

2 The model

We consider a population of haploid individuals whose genome is represented as a string of L bits. Each bit represents a *locus* and the Boolean values it can take are regarded as alternative allelic forms. In particular the value 0 refers to the *wild-type* allele while the value 1 to the least deleterious mutant. The phenotype, in agreement with the theory of quantitative traits [19, 20] is just the sum of these bits. The mutation is simply implemented by flipping a randomly chosen bit from 0 to 1 or *vice versa*. This kind of mutations can only turn a phenotype x into one of its neighbors $x+1$ or $x-1$ and they are therefore referred to as *short range mutation*.

The model is basically composed by a selection step followed by a reproduction procedure. The population grows according to a logistic law with overlapping generations. The choice of the fitness landscape is of paramount importance in selection. In our model we consider a static and a dynamic component of fitness. The static component describes the adaptation to environmental factors not related to the population itself *e.g.* abiotic factors such as climate, temperature, etc. The dynamic component describes how the interactions with other members of the population (competition, predation, mutualism) affect the fitness and it changes in time as a function of the population itself. In our very simplified model we considered only competition. The static component of the fitness is defined as :

$$H_0(x) = e^{-\frac{1}{\beta}\left(\frac{x}{\Gamma}\right)^{\beta}}$$

Since the individuals with similar phenotypes are those sharing the largest quantity of resources , the competition is the stronger the more similar their phenotypes are. This is why we chose an exponentially decreasing competition kernel so that the complete expression of fitness in our model is given by:

$$H(x) = H_0(x) - J\sum_y e^{-\frac{1}{\alpha}\left|\frac{x-y}{R}\right|^{\alpha}} P(y) \qquad (1)$$

It is now easy to work out the fitness function that accounts for the fact that the number of offsprings must always be a positive quantity:$A(x) = e^{H(x)}$. The number of survivors of phenotype i after selection is : $n'(i) = n(i)(1-m)\frac{A_i}{A}$ where m is the fitness-independent mortality and $\left(1-\frac{A_i}{A}\right)$ is the fitness-dependent mortality. The selection criterion is thus based on the very simple idea that individuals with higher than average fitness have the highest chance to survive.

After selection, the population goes through the reproduction process. The mechanism of reproduction of non recombinant individuals is very simple and reminds that of virus. A copy of the genome is

made and a mutation is introduced on a random bit according to a mutation rate μ.

The reproduction of recombinants is slightly more complex. First, the genome of the offspring is built by choosing for each *locus* the allele of the first or second parent with the same probability and then a mutation is introduced with the same procedure employed in the case of non recombinants. In our model we therefore assume absence of *linkage* but it must be remembered that it is reasonable only in the case of very long genomes subdivided into many independent chromosomes.

The assortativity is introduced through a parameter Δ which represent the maximal phenotypic distance still compatible with reproduction. In other words, if the first parent has phenotype i its partners must be chosen in the range $[i-\Delta, i+\Delta]$.

In order to speed-up simulations, we used a simulated annealing technique: the mutation rate μ depends on time as

$$\mu(t) = \frac{\mu_0 - \mu_\infty}{2}\left(1 - \tanh\left(\frac{t-\tau}{\delta}\right)\right) + \mu_\infty,$$

which roughly corresponds to keeping $\mu = \mu_0$ up to a time $\tau - \delta$, then decrease it linearly up to the desired value μ_∞ in a time interval 2δ and continue with this value for the rest of simulation. Simulations show that that the variance and mean reach their asymptotic value very quickly.

3 Simulations

If competition is absent and the static fitness landscape is steep, the distribution of both recombinants and non recombinants (with initial binomial distributions centered at $x = 7$) move towards the fitness maximum at $x = 0$. The high mutation rate ($\mu_0 = 10^{-4}$, $\mu_\infty = 10^{-6}$) provides non recombinants with a velocity of movement in the phenotypic space comparable to that of recombinants. As a consequence the non recombinants usually reach the $x = 0$ position earlier than recombinants that are thus led to extinction. In some runs however, it may happen that recombinants and non recombinants reach the $x = 0$ position in the same time or that recombinants arrive a little earlier than non recombinants. In any case, the recombinants do not have enough time to establish a large colony and when the non recombinants reach $x = 0$, thanks to their higher fertility, they soon overwhelm recombinants that tend to disappear.

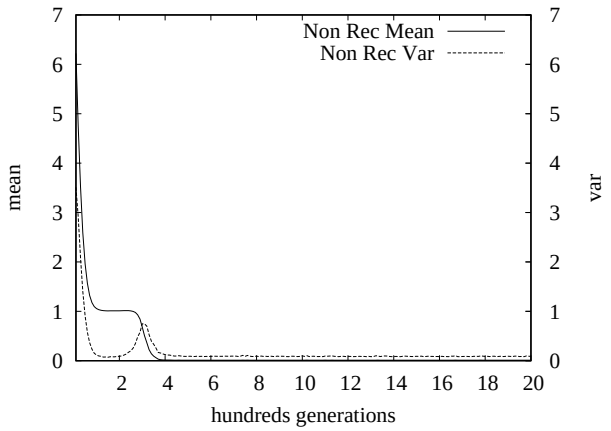

Figure 1: Steep static fitness ($\beta = 1$, $\Gamma = 14$) and absence of competition ($J = 0$): mean and variance of non recombinant distribution. Mating range: $\Delta = 14$. Initial frequency of recombinants: 0.1; initial distribution parameters: $p = q = 0.5$, initial population size $N_0 = 1000$, carrying capacity $K = 10000$. Annealing parameters: $\mu_0 = 10^{-4}$, $\mu_\infty = 10^{-6}$, $\tau = 1000$, $\delta = 100$. Total evolution time: 10000 generations; for the sake of clarity only 2000 and 3000 generations were displayed in the plots of mean and variance of non recombinants and in the frequency plots respectively.

In Figure 1 and 2 we report a simulation in which recombinants reach $x = 0$ first. The mean of non recombinants decreases in a step-like way as they establish a delta-peak in $x = 1$ that later moves to $x = 0$; their variance conversely decreases abruptly from a very high value related to the wide initial distribution to almost zero forming a little hump in correspondence to the shift from the peak in $x = 1$ to that in $x = 0$. The variance and mean of recombinants conversely, first decrease when they establish a peak in $x = 0$ and then increase again when the peak in $x = 0$ disappears and the few survivors become distributed according to a wide and flat distribution.

The simulation illustrated in Figures 1 and 2 shows that when the initial mutation rate is as high as 10^{-4} the non recombinants are characterized by a high velocity of movement in the phenotypic space so that they are able to reach the $x = 0$ position before or in the same time as the recombinants. On the other hand, if we set $\mu_0 = 10^{-5}$ the mobility of

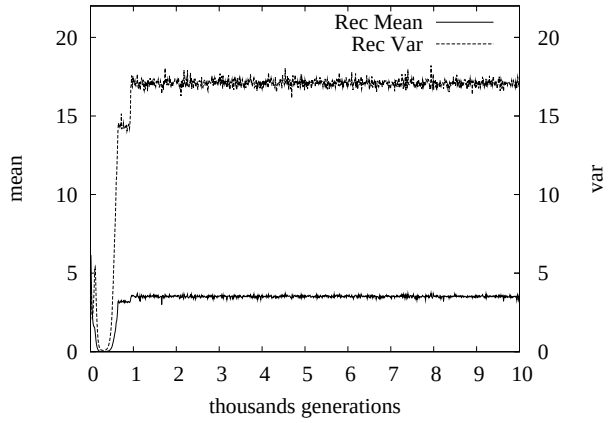

Figure 2: Steep static fitness ($\beta = 1$, $\Gamma = 14$) and absence of competition ($J = 0$): mean and variance of recombinant distribution. Mating range: $\Delta = 14$. Parameters as in Figure 1

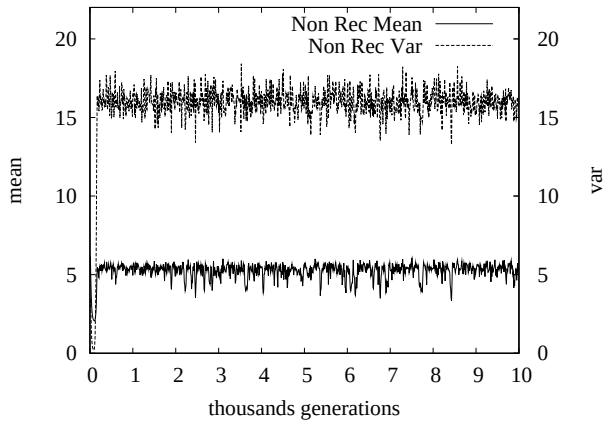

Figure 3: Steep static fitness ($\beta = 1$, $\Gamma = 14$) and absence of competition ($J = 0$) in a low mutation regime: mean and variance of non recombinant distribution. Mating range: $\Delta = 14$. Initial frequency of recombinants: 0.1; initial distribution parameters: $p = q = 0.5$, initial population size $N_0 = 1000$, carrying capacity $K = 10000$. Annealing parameters: $\mu_0 = 10^{-5}$, $\mu_\infty = 10^{-6}$, $\tau = 1000$, $\delta = 100$. Total evolution time: 10000 generations.

non recombinants is significantly reduced as compared to that of recombinants that can rely on the recombination mechanism. As a consequence, the recombinants can reach the $x = 0$ position much earlier than non recombinants and are therefore able to establish there large colonies that, due to sampling effects, cause the extinction of the peaks of non recombinants in $x = 1$ or $x = 2$.

In Figures 3 and 4 we display the results of a typical run. The mean and variance of non recombinants decrease abruptly when they form a peak in $x = 2$. When this peak becomes extinct however, the non recombinant distribution becomes extremely wide and flat, with mean and variance showing wide fluctuations because of random sampling effects in a small subpopulation. The variance and mean of recombinants on the other hand decrease monotonically as the distribution moves towards $x = 0$ and in the same time narrows down until it becomes a delta-peak.

Let us consider an example of the case $M = 0.5$, $p = q = 0.5$, *i.e.* recombinants and non recombinants have the same initial frequency and the initial distributions are both centered in the middle of the phenotypic space. In a regime of weak competition such as $J = 0.8$ both distributions move towards the maximum of static fitness but the $x = 0$ position is first reached by the recombinants. As the peak of recombinants in $x = 0$ becomes more and more populated it exerts a stronger and stronger competition on the peak of non recombinants in $x = 2$ (or $x = 1$)

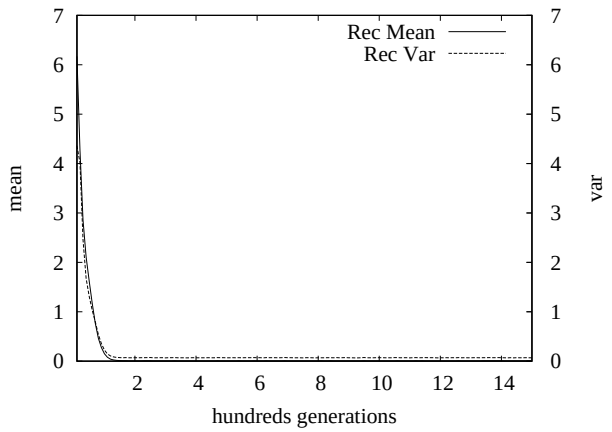

Figure 4: Steep static fitness ($\beta = 1$, $\Gamma = 14$) and absence of competition ($J = 0$) in a low mutation regime: mean and variance of recombinant distribution. Mating range: $\Delta = 14$. Initial frequency of recombinants: 0.1; initial distribution parameters: $p = q = 0.5$, initial population size $N_0 = 1000$, carrying capacity $K = 10000$. Annealing parameters: $\mu_0 = 10^{-5}$, $\mu_\infty = 10^{-6}$, $\tau = 1000$, $\delta = 100$. Total evolution time: 10000 generations; for the sake of clarity only 1500 generations were displayed.

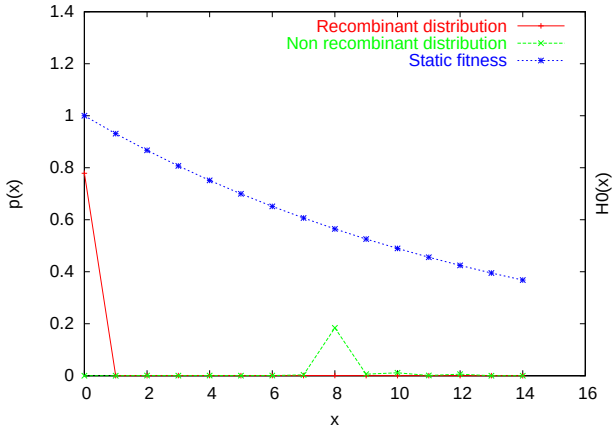

Figure 5: Steep static fitness ($\beta = 1$, $\Gamma = 14$) and weak competition ($J = 0.8$) in a random mating regime ($\Delta = 14$): final distribution. Initial frequency of recombinants: 0.5; initial distribution parameters: $p = q = 0.5$, initial population size $N_0 = 1000$, carrying capacity $K = 10000$. Annealing parameters: $\mu_0 = 10^{-5}$, $\mu_\infty = 10^{-6}$, $\tau = 1000$, $\delta = 100$. Total evolution time: 10000 generations.

which is therefore led to extinction. The disappearance of the non recombinant peak in $x = 2$ relieves competition in the middle of the phenotypic space therefore allowing the appearance of a new peak of non recombinants in $x = 7$ or $x = 8$. The central peak of non recombinants is stabilized by competition: the central region of the phenotypic space in fact becomes a favourable position because the low value of the static fitness is compensated by the low level of competition experienced in that location. In Figure 5 we show the plots of the final distribution in a typical case.

4 Conclusions

A microscopic model was developed for the study of the evolution of recombination in an environment whose features depend on the population itself (the frequency distribution of phenotypes determines competition and hence fitness). The simulations show, that under these conditions recombination provides organisms with a high velocity of movement in the phenotypic space, so that recombinants reach the regions of the phenotypic space with the highest fitness much earlier than non recombinants whose ability of movement is limited by the low mutation rate. The simulations also show (in agreement with the findings in our study on sympatric speciation [21]) that competition acts as a stabilizing force allowing the survival of strains in regions of the phenotypic space where the low static fitness level is counteracted by a low competition pressure.

References:

[1] R. Butlin. The costs and benefits of sex: new insights from old asexual lineages. *Nature Reviews Genetics*, 3:311–317, 2002.

[2] J. Maynard Smith and E. Szathmàry. *The Major Transitions in Evolution*. Freeman/Spectrum, Oxford, 1995.

[3] R.E. Michod. Genetic error, sex and diploidy. *J. Hered.*, 84:360–371, 1993.

[4] R.J. Redfield. Do bacteria have sex ? *National reviews of Genetics*, 2:634–639, 2001.

[5] T. Cavalier-Smith. The phagotrophic origin of eukaryotes and phylogenetic classification of protozoa. *International Journal of Systematic Evolutionary Microbiology*, 52:297–354, 2002.

[6] J.R. Peck. A ruby in the rubbish: beneficial mutations, deleterious mutations and the evolution of sex. *Genetics*, 137:597–606, 1994.

[7] D. Charlesworth, B. Charlesworth, and M.T. Morgan. The pattern of neutral molecular variation under the background selection model. *Genetics*, 141:1619–1632, 1995.

[8] N.H. Barton. Linkage and the limits of natural selection. *Genetics*, 140:821–841, 1995.

[9] J.T. Manning and D.J. Thompson. Muller's ratchet and the accumulation of favorable mutations. *Acta Biotheor.*, 33:219–225, 1984.

[10] M. Kimura. On the probability of fixation of mutant genes in a population. *Genetics*, 47:713–719, 1962.

[11] B. Charlesworth. The effect of background selection against deleterious mutations on weakly

selected, linked variants. *Genet. Res.*, 63:213–227, 1994.

[12] R.A. Fisher. *The genetical theory of natural selection*. Oxford University Press, 1930.

[13] H.J. Muller. Some genetic aspects of sex. *American Naturalist*, 66:118–138, 1932.

[14] J.F. Crow and M. Kimura. Evolution in sexual and asexual populations. *American Naturalist*, 99:439–450, 1965.

[15] W.R. Rice. Experimental tests of the adaptive significance of sexual recombination. *Nature Reviews Genetics*, 3:241–251, 2002.

[16] S.P. Otto and T. Lenormand. Resolving the paradox of sex and recombination. *Narure Reviews*, 3:252–261, 2002.

[17] N.H. Barton. A general model for the evolution of recombination. *Genet. res.*, 65:123–144, 1995.

[18] G. Bell. *The masterpiece of nature: the evolution and genetics of sexuality*. University of california press, Berkley, 1982.

[19] M. Lynch and B. Walsh. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates Inc., 1998.

[20] D.S. Falconer and T.F.C. Mackay. *Introduction to Quantitative Genetics*. Addison-Wesley Publishing Company, 1996, 4th Edition.

[21] F. Bagnoli and C. Guardiani. A model of sympatric speciation through assortative mating. Submitted to Physical Review A.