International Journal of Modern Physics C
Vol. 18, No. 8 (2007) 1329–1338

World Scientific
www.worldscientific.com

INBREEDING AND OUTBREEDING DEPRESSIONS IN THE PENNA MODEL AS A RESULT OF CROSSOVER FREQUENCY

K. BOŃKOWSKA, M. KULA, S. CEBRAT and D. STAUFFER*

Department of Genomics, Wrocław University
ul. Przybyszewskiego 63/77, 51-148 Wrocław, Poland
** stauffer@thp.uni_koeln.de*

Received 23 February 2007
Accepted 26 February 2007

The population in the sexual Penna aging model is first separated into several reproductively isolated groups. Then, after equilibration, sexual mixing between the groups is allowed. We study the changes in the population size due to this mixing and interpret them through a counterplay of purifying selection and of haplotype complementarity.

Keywords: Monte Carlo simulation; sexual reproduction; crossover; Penna ageing model.

1. Introduction

In sexual reproduction, as opposed to asexual reproduction, the genomes of the two parents are mixed, and within the diploid genome of each parent happens crossover. This way of reproduction has advantages as well as disadvantages compared with asexual cloning of haploid genomes. An advantage is that bad recessive mutations do not affect the health if they are present in only one of the two haplotypes (= sets of genetic information). A disadvantage is the reduced number of births if only the females produce offspring while the males consume as much food and space as the females. Moreover, crossover of two different genomes may produce a mixture which is fitter than each of the two parents but also the one which is less fit, as seen in the extreme case for donkeys and horses (outbreeding depression). For small populations, the probability is higher that the two parents have the same bad recessive mutation which, therefore, diminishes the health of the individual (inbreeding depression).

2. Standard Model

We try to simulate these effects in the standard sexual Penna aging model, at first with a length $L = 32$ for each of the two bit-strings representing the genome,

*Visiting from Institute for Theoretical Physics, Cologne University, D-50923 Köln, Euroland.

a minimum reproduction age $R = 8$, six dominant loci among the 32 total loci, one new random irreversible heritable mutation per bit-string at birth, as well as a lethal threshold T for the number of active detrimental mutations, and B births per mature female and iteration. We deviate from the published programs[1] as follows: The Verhulst factor, a death probability $N/N_{\max}$ at population N with carrying capacity $N_{\max} = 10^3$, 10^4, 10^5 due to limited food and space, was applied to the birth only and not to adults; the initial age distribution was taken as random when needed; the birth rate B was reduced from 4 to 1, T from 3 to 1 (that means a single active mutation kills), and mostly only 10^4 instead of 2×10^4 time steps were made. (One time step or iteration is one Monte Carlo step for each individual.) Furthermore, the whole population was for most of the simulated time separated into G different groups such that females look for male partners only within their own group, with a separate Verhulst factor applying to each group. For the last $\Delta \ll 10^4$ time steps this separation into groups was dissolved: Then, females could select any male, and only one overall Verhulst factor is applied to the whole population. Finally, the crossover process within each parent before each birth was not made always but only with a crossover probability r.

If there would be no inbreeding depression then during the first longer part of the simulation, the total number N_1 of individuals would be independent of the number G of groups into which it is divided. And if then there are no advantages or disadvantages of outbreeding, the population N_2 during the second shorter part, $10^4 - \Delta < t < 10^4$, would be the same as the preceding population N_1 during the last section, $10^4 - 2\Delta < t < 10^4 - \Delta$, of the longer first part. We will present the data which shows that this is not the case. Similar simulations for $G = 2$ groups were published earlier.[2]

A difficulty in such simulations is the Eve effect: After a time proportional to the population size, everybody has the same female (Eve) as ancestor, with all other offspring having gotten less fit genomes due to random mutations, or just less luck, and thus having died out. (Similarly there is also one common Adam, but these two individuals in general never met.) If we would divide the whole population into many groups without further changes, the Eve effect would let all groups but one die out and thus destroy the separation. Therefore, for the first long period of separation we used separate Verhulst factors for each group, stabilizing its population, while for the second shorter part of mixing we used mostly $\Delta = N_{\max}/100$.

Figure 1 shows for $N_{\max} = 10^4$ the dependence on the crossover probability for the populations N_1 before and N_2 after mixing. We see that the mixing always increases the population, that means one has no outbreeding depression but an outbreeding advantage. Figure 2 confirms this advantage but also shows the inbreeding depression: The larger the number G of groups (and thus the smaller the group size) is, the smaller are the two populations N_1 and N_2. (The difference between N_1 and N_2 fluctuates less than these numbers themselves since N_2 is strongly correlated with N_1. For longer observation times 10^5 and 10^6 some of the points to the right go to zero.) Also, for the larger population in Fig. 2, the number of

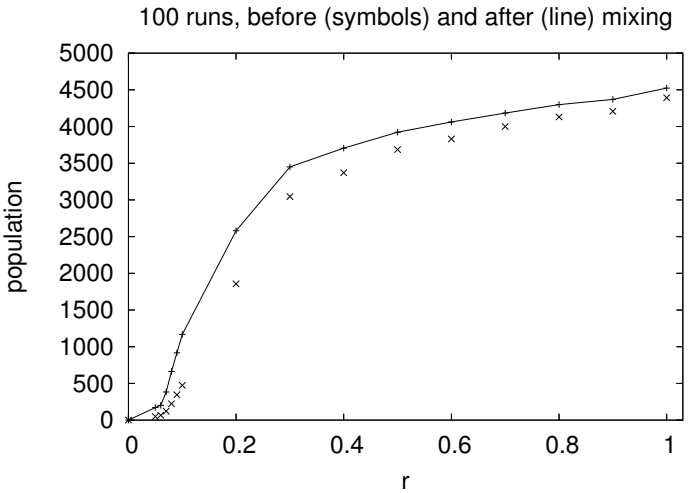

Fig. 1. Average over 100 simulations with $G = 10$ groups each, $\Delta = 100$.

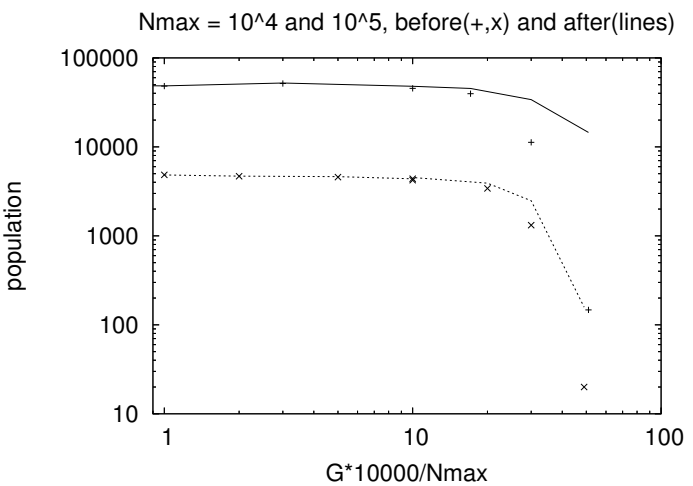

Fig. 2. Average over 10 simulations with a large (top) and 100 with a small (bottom) population, versus number G of groups; $r = 1$, $\Delta = 1000$ and 100, respectively. For the larger population G is divided by 10 so that data for the same number of individuals per group have the same horizontal coordinate. We see nice scaling.

groups can be larger before the population becomes extinct. Figure 3 shows the time dependence of the outbreeding effect by mixing between groups allowed after 9900 (part a) and 9000 (part b) time steps. Figure 3(a) shows summed populations from 100 simulations with a small population ($G = 10$) and 10 simulations of a large population ($G = 100$) versus time after mixing started; $r = 1$ in both cases. For much larger populations of 5 million and still $G = 10$, no such effect of mixing is seen. Part b shows for the high reproduction age R of the following figures one

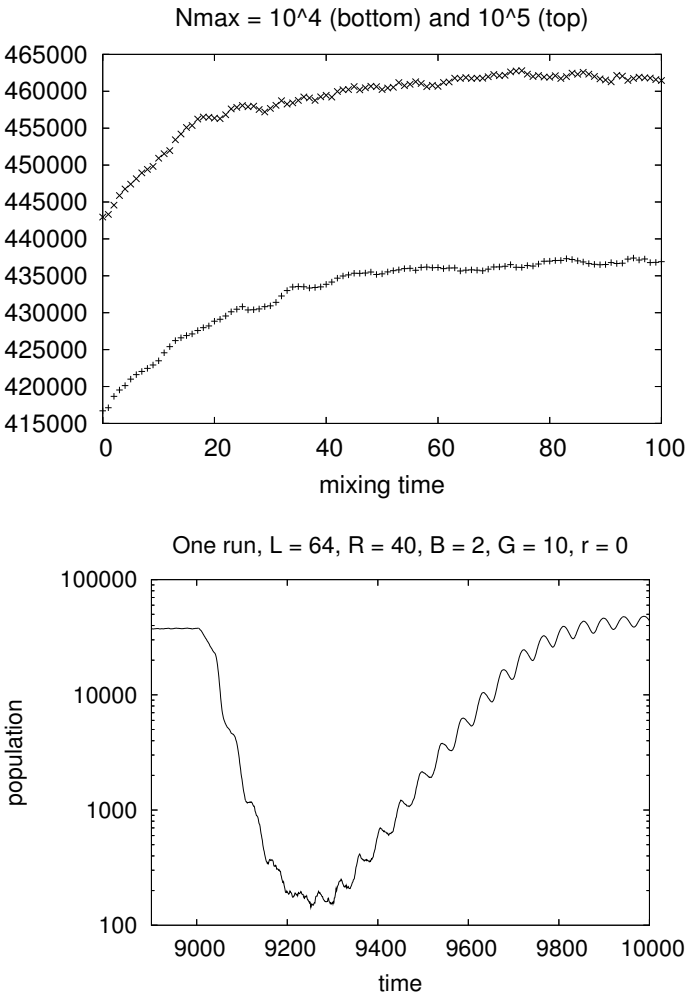

Fig. 3. Time dependence of outbreeding advantage (part a) and outbreeding depression (part b).

example of the outbreeding depression (bottleneck[3]) followed by a recovery with oscillations of period R after mixing was allowed from time 9001 on; $N_{\max} = 10^5$.

We also checked for the influence of r in the case when the minimum age of reproduction R is 5/8 of the length L of the bit-strings, i.e., larger than the value of 8 used before, and when L is different from 32 used in Figs. 1–3(a). In these simulations we also assumed all mutations to be recessive, in contrast to the 6 out of 32 dominant bit positions for Figs. 1–3(b). Figure 4 shows for $L = 32$ and a birth rate $B = 4$ a minimum of the population at intermediate r for one group, and for 50 groups a monotonic behavior but with outbreeding depression at small r and outbreeding advantage at big r. This population minimum is seen for $L = 64$ and

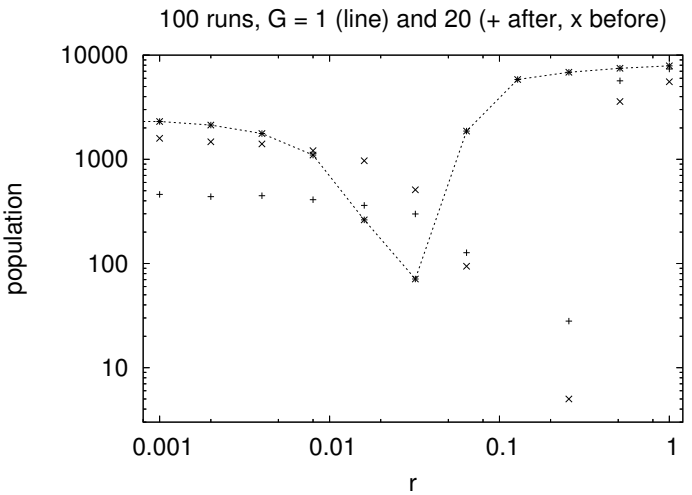

Fig. 4. Average over 100 simulations with a small population, high minimum age of reproduction, and $G = 1$ and 20. For $G = 1$ there is always complete mixing. Note the double-logarithmic scales, also in Figs. 5 and 6.

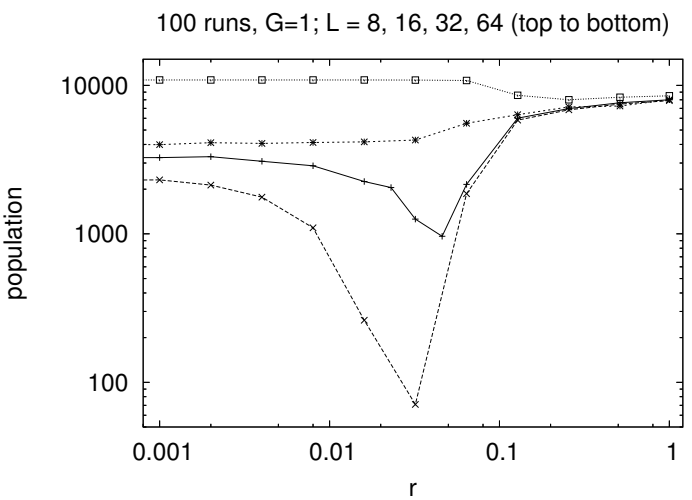

Fig. 5. Average over 100 simulations with a small population, high minimum age of reproduction and various lengths L of the bit-strings, using a birth rate $B = 128/L$.

32 but not for 16 (Fig. 5). Figure 6 shows the dependence on population size. (Our data before and after mixing are averaged over $\Delta = 100$ or 1000 iterations. When outbreeding depression occurs it may happen that later the population recovers: Fig. 3(b).)

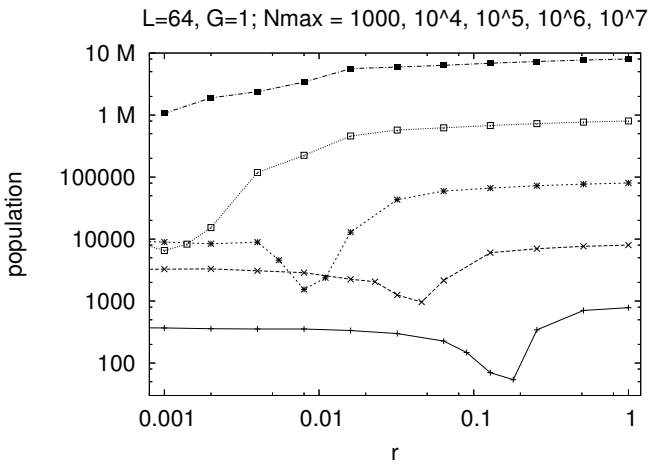

Fig. 6. Dependence on population size for $N_{\max} = 10^3, \ldots, 10^7$, averaged over 1000 to one sample. $L = 64$, $B = 2$, $t = 10^4$. Near the minimum, equilibrium has not yet been reached; populations still decrease there.

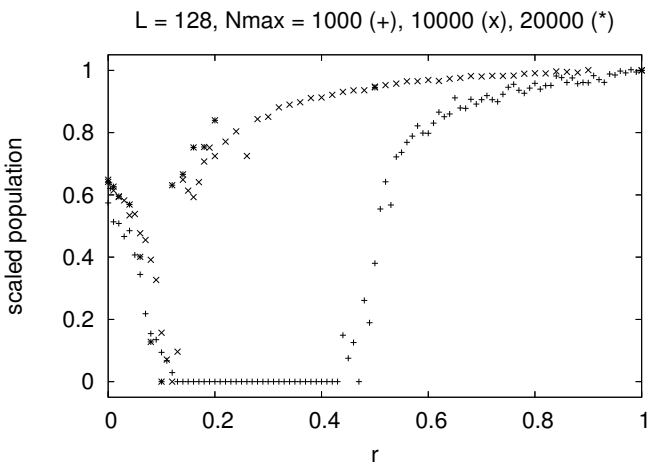

Fig. 7. Relation between normalized population size and the crossover rate. The population size was divided by the population evolving with crossover rate $r = 1$.

3. Interpretation

To study the inbreeding and outbreeding depressions in detail we have analyzed the results of the simulations of single population of different size under different regime of intragenomic recombinations (crossover rate r). Parameters for these simulations have been slightly changed to get better results: $L = 128$, $R = 80$, $N_{\max} = 1000$–$20\,000$, crossover probability $r = 0$–1, $B = 1$, time of simulations $= 5{\times}10^5$ iterations. In Fig. 7, the relation between the size of population and the crossover probability for three different environment capacities are shown.

Populations in the smallest environment ($N_{\max} = 1000$) survive with $r = 0$ but their sizes decrease with increasing r and are extinct for r set between 0.12 and 0.4. Under larger crossover rates populations survive and their sizes are larger than those obtained for $r = 0$ (see plots in Fig. 7 where sizes of populations were normalized by the size of population under $r = 1$). Larger populations ($N_{\max} = 10\,000$) are extinct in a very narrow range of crossover rates close to 0.12, and populations with $N_{\max} = 20\,000$ become extinct at slightly lower crossover rates. Nevertheless, all populations have larger sizes when the crossover rate is of the order of 1 per gamete production (the highest tested).

This nonlinear relation between the size of population and the crossover rate could be explained on the basis of the genetic structure of the individual genomes in the simulated populations. In Fig. 8 we have shown the frequency of defective genes in the genetic pool of populations for $N_{\max} = 10\,000$ under crossover rates 0, 0.1, and 1. The frequency of defective genes expressed before minimum reproduction age ($R = 80$) in populations without crossover is 0.5. Since $T = 1$, if the distribution of defects would be random, the probability of any individual to survive until the reproduction age R would be 0.75^R (negligibly small for large $R > 30$). Thus, to survive, individuals have to complete their genomes of two complementing bit-strings (haplotypes). For more efficient reproduction, the number of different haplotypes should be restricted and in fact there are only two different complementing haplotypes in the whole population as it was shown in Ref. 4. In such populations, the probability of forming the offspring surviving until the reproduction age is 0.5. Note that recombination at any point inside the part of the genome expressed before reproduction age R produces a gamete which is not complementary to any other gamete produced without recombination or with recombination in any other point.

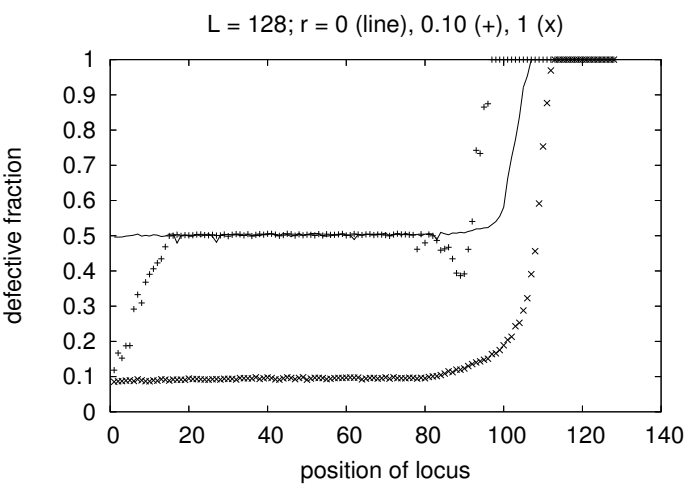

Fig. 8. Distribution of defective genes in the genomes of populations evolving under different crossover frequencies.

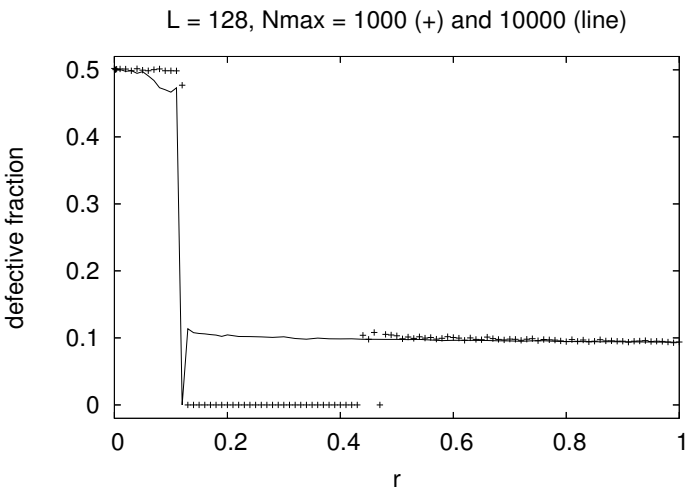

Fig. 9. Relation between crossover frequency and average frequency of defective genes in the sections of genomes expressed before the reproduction age R. Note, fractions equal 0 mean that the populations with the given crossover frequency died out.

Thus, crossovers in such populations are deleterious for the offspring. On the other extreme, with crossover probability = 1, populations are under purifying selection. The fraction of defective genes in the population is kept low (about 0.1, compared with 0.5 without recombination), to enable the surviving of the offspring until their reproduction period. The critical crossover frequency close to 0.12 is connected with a sharp transition from these two strategies of genomic evolution: complementarity and purifying selection. In Fig. 9, the frequency of defective genes expressed before the reproduction age is plotted. For lower crossover rates the fractions of defective genes are kept at the level of 0.5, for higher crossover rates they are close to 0.1. Close to the critical frequency of crossover, defective genes located at both ends of the region of genomes expressed before the reproduction age are forced to obey the purifying selection which eliminates some defects (Fig. 8).

In the case of small populations, the probability of meeting two closely related partners (high inbreeding coefficient) is high and as a consequence, there is a higher probability of meeting two defective alleles in the same locus in the zygote which determines a phenotypic defect and eliminates the offspring from the population. In such condition, the strategy of completing the genome of two complementing haplotypes is more effective. Nevertheless, this strategy is not the best if effective populations are very large, with low inbreeding coefficient, when the probability of meeting two identical haplotypes is negligible. Thus, comparing very large populations with very small ones we can observe the inbreeding depression. On the other hand, this strategy in small populations leads to the emerging of a very limited number of different haplotypes in the populations (in extreme only two). These haplotypes are characterized by a specific sequence of defective alleles. Independent

simulations generate haplotypes with different sequence of defective alleles. Mixing two or more populations evolving independently decreases the probability of meeting in one zygote two complementing haplotypes, this difference results in outbreeding depression (seen in Figs. 3(b) and 4).

This purifying selection is nothing but Darwin's selection of the fittest and dominates for large r. Evidence for our new claim of complementary haplotypes at small r is shown in Fig. 10, with the distribution of Hamming distances between the two haplotypes of one individual. (The Hamming distance is the number of bits which differ in a position-by-position comparison of the two bit-strings. Since at old age all bits are set, Fig. 10 analyzes only the first $R = 40$ of the $L = 64$ bits. Only the two bit-strings of the same individual are compared, not those of the two parents at the birth of their child.) For $r = 0.128$ we see a single peak at intermediate Hamming distances, while for $r = 0$ the distribution splits into a smaller part near zero (nearly identical haplotypes) and a larger part near the maximum of 40 (nearly complementary haplotypes). This picture is confirmed by the probability for a defective bit of an age up to R, which is 0.5 for $r = 0$ and about 0.1 for $r = 0.128$, and by ordering the final population into sets of identical bit-strings (not shown). Thus, only in the $r = 1$ case does selection try to remove deleterious mutations, while for $r = 0$ it splits the whole set of bit-strings into two sets A and B such that a B haplotype is nearly the complement of an A haplotype. Since even for $r = 0$ we combine for each birth one maternal bit-string with one paternal bit-string of another individual, babies may also get two A or two B bit-strings, are then less fit for survival and form the leftmost part of the distribution in Fig. 10.

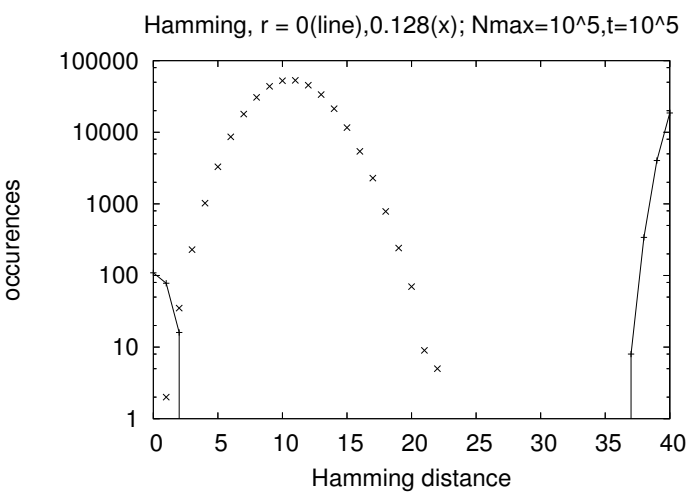

Fig. 10. Distribution of Hamming distances for small (one peak) and large (two peaks) crossover rates; $L = 64$, $G = 1$.

4. Conclusion

We varied the parameters of the sexual Penna aging model, in particular by separating the population into reproductively isolated groups and/or having longer bit-strings and a high minimum age of reproduction. We could observe and interpret inbreeding depression, outbreeding depression, and outbreeding advantage, through the counterplay of purifying selection and of haplotype complementarity. Purifying selection tries to have as few mutations in the bit-strings, like haplotype 00000000 for $L = 8$, while haplotypes 01100101 and 10011010 are complementary. In both cases, deleterious effects from mutations are minimized.

Note Added in Proof

Our results in Figs. 4 and 6 depend strongly on the mutation rate if it is varied between 1/2 and 1.

References

1. S. M. de Oliveira, P. M. C. de Oliveira and D. Stauffer, *Evolution, Money, War and Computers* (Stuttgart and Leipzig, Teubner, 1999); D. Stauffer, S. M. de Oliveira, P. M. C. de Oliveira and J. S. Sá Martins, *Biology, Sociology, Geology by Computational Physicists* (Elsevier, Amsterdam, 2006).
2. A. Łaszkiewicz, Sz. Szymczak and S. Cebrat, Speciation effects in the Penna aging model, *Int. J. Mod. Phys. C* **14**, 765–774 (2003); see also A. O. Sousa, S. M. de Oliveira and A. T. Bernardes, Simulating inbreeding depression through the mutation accumulation theory, *Physica A* **278**, 563–570 (2000).
3. K. Malarz and D. Stauffer, Search for bottleneck effects in Penna ageing and Schulze language model, q-bio.PE/0609051 at arXiv.org, to appear in *Adv. Complex Syst.*
4. M. Zawierta, P. Biecek, W. Waga and S. Cebrat, The role of intragenomic recombination rate in the evolution of population's genetic pool, *Theory in Biosciences* **25**, 123 (2007).