

WHY SEX? — MONTE CARLO SIMULATIONS OF SURVIVAL AFTER CATASTROPHES

J. S. SÁ MARTINS* and S. MOSS DE OLIVEIRA†

*Instituto de Física, Universidade Federal Fluminense
Av. Litorânea s/n, Boa Viagem, Niterói 24210-340, R.J, Brazil*

**E-mail: jssm@if.uff.br*

†E-mail: suzana@if.uff.br

Received 9 March 1998

Revised 11 March 1998

Using the Penna bit-string model for biological ageing we compare two kinds of reproductive regimes: Sexual reproduction (SR) and meiotic parthenogenesis (MP). The last one is a common type of asexual reproduction *with recombination*, found in diploid organisms. We show that although both regimes present roughly the same survival rates, the diversity generated by SR is much larger, and can prevent the extinction of a population submitted to a natural disaster. The fixation of bad genes inside an MP population, after many generations, explains our results. We also study the consequences of cloning (simple copy) on population diversity.

Keywords:

1. Introduction

The Penna bit-string model for biological ageing was published in 1995,¹ and since then around 40 papers have been published using this model to study different characteristics of real populations. The catastrophic senescence of Pacific Salmon is an example,² and many others can be found in Ref. 3. In this paper, the idea has been to use the model in order to compare sexual reproduction with the parthenogenetic asexual one.

The question of why sex evolved is a very old one, and remains unsolved. It is a well-known fact that simple asexual reproduction (without recombination, for haploid organisms) is a very fast process, but it results in a poor genetic diversity, in contrast to sexual reproduction, which generates a higher diversity but a lower number of offspring (besides the risk of dilution of well-fitted genomes). The meiotic parthenogenetic reproductive regime is somewhere between simple asexual reproduction and the sexual one: The organism is diploid (with a pair of homologous chromosomes, as in the sexual case), but the recombination process, to produce the offspring, occurs between the homologous chromosomes of *a single* individual. In this way the MP regime avoids the higher accumulation of deleterious mutations

presented by simple asexual reproduction (AR), and does not present the high risk of genome dilution that may occur in the SR.

Using the Penna model for biological ageing, Stauffer *et al.* have already compared the survival rates of AR and SR,⁴ showing that SR presents a higher survival rate. However, using the same approach, Bernardes⁵ has shown that the survival rates of MP and SR are equivalent, and has found no evidence of advantage of sex in large populations except when male competition is introduced. We study exactly the same problem as Bernardes did, but compare the resulting diversities instead of the survival rates. We followed Dasgupta⁶ in promoting a catastrophe in order to verify and measure the advantage of a higher diversity, which turns out to be the advantage of sex. Since the Penna model uses the genetic code (genome) to represent each individual, it allows one to study how the genetic pattern of a population evolves. In so doing, it is possible to relate the fixation of bad genes with the possibility of survival if an ecological disaster occurs. It is also possible to verify how the genetic diversity of an already stable sexual population is affected by the introduction of clones.

This paper is organized in the following way: In Sec. 2 we explain the reproductive regimes within the Penna model, and how the genetic diversity is measured; in Sec. 3 we present our comparisons and results, and in Sec. 4, our conclusions.

2. Reproductive Regimes and Diversity

In this paper we are interested in studying, within the Penna model, the reproductive regimes of *diploid* organisms, that is, organisms that present an homologous pair of chromosomes. In this case, the genome of each individual is represented by *two* bit-strings of zeroes and ones, *that are read in parallel*. Each pair of homologous bits (genes) correspond to one “year” in the individual lifetime. The first bit of one string together with the first bit of the other string contain the information of what will happen to the individual in its first year of life. The second bits contain the information about the second year of life, and so on. Since we have been working with strings of 32 bits, each individual can live at most for 32 years. In these strings genetic diseases are represented by 1 bit. If an individual has, for instance, both of the third bits (one of each string) equal to one, it means that the individual will start to suffer from the effects of an inherited disease from age three until death. If the same thing happens at another position, the individual will accumulate one more disease. There is a limit T for the number of such diseases; at every time step (“year”) a new pair of homologous bits is read and the number of diseases *until that year* is computed. If it becomes equal to T the individual dies. In the example given above the individual is an homozygotous one, that is, with 1 bit at the third position of *both strings*. In this case the individual will surely get sick when it is three years old. In the opposite case, when the individual inherits zero bits in the same positions of each string, it does not get sick. When the individual is an heterozygotous one, with different bits at the same position of each string, the

possibility of sickness depends on whether that position is *dominant* or not. We define at the beginning of the simulation the number of dominant positions and randomly choose them. These positions hold for all the individuals, and remain fixed during the whole simulation.⁴

Besides dying if the number of allowed diseases is exceeded or when reaching 33 years, an individual may also die for lack of space and food. Such a case is taken into account through the Verhulst factor $V = 1 - N(t)/N_{\max}$, where $N(t)$ is the actual size of the population at time t and N_{\max} is the maximum size (carrying capacity) the population can sustain, which is defined at the beginning of the simulation. At every time step and for each individual, a random number between 0 and 1 is generated and compared with V . If this random number is greater than V , the individual dies independently of its age or genome.

All the rules mentioned above hold for both MP and SR. The differences between them appear at the moment of reproduction, that is, in the way that the offspring genome is constructed. Let us define firstly the sexual reproduction. In this case half of the initial population $N(0)$ corresponds to males and the other half to females, and there is a minimum reproduction age R that holds for both sexes. Every female that reaches the age R , randomly chooses a male with age equal or greater than R to mate, generating b offspring. The mating process is repeated every year until death. The offspring genome is constructed in the following way: the mother genome, for instance, is cut in a random position (crossing), generating 4 bitstring pieces (see Fig. 1). Two complementary pieces, each one coming from one of the original strings, are joined (recombination) to form the offspring string that contains the genetic charge of the mother. After this, M random *deleterious* mutations are included. The same procedure is repeated with the father genome, to produce the

sexual reprod.		meiotic par.	asexual
mother	father	crossing	parent
crossing	crossing	10 0100	101010
00 1001	100 001	10 0110	
11 0011	110 011	recombination	
recombinations		100110	
111001	100011	copy	copy
mutations		100110	101010
111011	110011	mutations	
↑	↑	↓	mutations
	111011	100111	111011
	110011	110110	↑
		↑	↑

Fig. 1. How the offspring genome is constructed for the three reproductive regimes.

second string of the baby. The sex of the baby is then randomly chosen. When only deleterious mutations are considered, and this is our case, whenever a 1 bit is randomly chosen in the parent genome, it remains equal to 1 in the offspring genome (no mutation occurs). However, if the randomly chosen bit of the parent genome is equal to zero, it is set to 1 in the offspring genome, corresponding to a bad mutation. We consider only bad mutations because they are around 100 times more frequent in Nature than the good ones⁷ (that would correspond to transforming a 1 bit of the parent into a zero bit of the offspring).

As mentioned in the abstract, meiotic pathenogenesis is an *asexual reproduction regime with recombination*, found in diploid organisms. In this case the simulation starts with $N(0)$ individuals, each one represented also by two bitstrings. When an individual succeeds in reaching age R , subjected to the same rules explained before, it produces b offspring. Following the example of Bernardes,⁵ the offspring genome is constructed by firstly cutting the two strings of the parent in a random position. Two complimentary pieces are chosen to be joined, exactly as in the sexual case (see Fig. 1), producing one new bitstring. Then *this string is copied*, and mutations are introduced. Due to the mutations, the two resulting strings are very similar, but not identical, in agreement with Stearns's explanation⁸ that meiotic parthenogenesis produces, in general, individuals that are homozygotous in all positions (loci).

As mentioned before, our results are related to the fixation of bad genes (1 bit) in the genomes, after many generations. In order to develop a better understanding of this fixation, which will be explained in the next section, we also studied the simple asexual reproduction of haploid organisms. In this case each genome is represented by a single bitstring of 32 bits, and whenever a 1 bit (disease) appears it is computed and the result compared with T . The offspring genome is a copy of the parent one, with M random mutations (see Fig. 1). There is also a minimum reproduction age R and the same dispute for space and food.

In order to verify the genetic diversity of the population we calculate the Hamming distance (number of different bits between the genomes) for all pairs of individuals. For the SR and MP cases, this comparison must be done in two steps. Suppose we have an individual i whose genome is represented by the bitstrings A_i and B_i , and an individual j with a genome given by A_j and B_j . Firstly we compare the strings A_i with A_j and B_i with B_j , storing the number of different bits between them as result 1. Secondly we compare A_i with B_j and A_j with B_i , storing the number of different bits as result 2. Then we compare result 1 with result 2 and keep only *the smaller one*. This procedure is necessary because the genome AB is completely equivalent to the genome BA . For the AR case only the first step is needed, since there is only one bitstring per genome.

Finally, we will also present the evolution of the genetic diversity when we introduce, in an already stable *sexual* population, a given number of *clones*. We consider the clones as copies of a single female (identical genomes). After being born, the clones (that are all females) are subject to exactly the same survival and reproduction rules as the whole population.

3. Results

All our simulations were run with the following parameters:

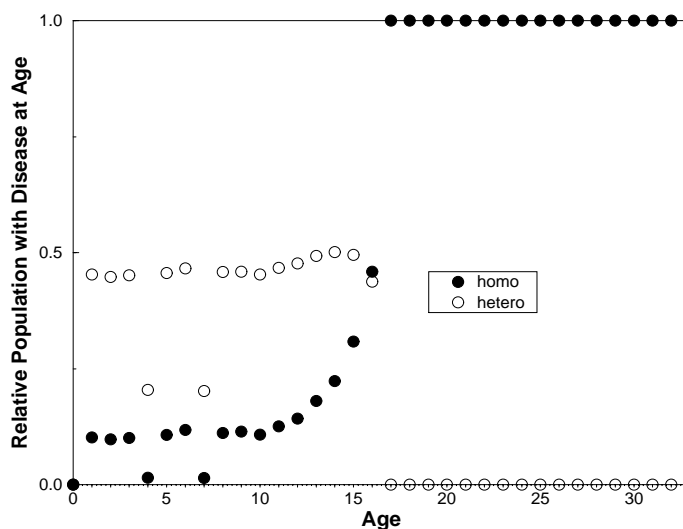
- (1) Maximum population size = 400 000;
- (2) Limit of allowed disease $T = 3$;
- (3) Minimum reproduction age $R = 10$;
- (4) Mutation rate $M = 2$;
- (5) Birth rate = 1.

Let us now present our results.

(a) Fixation of Bad Genes

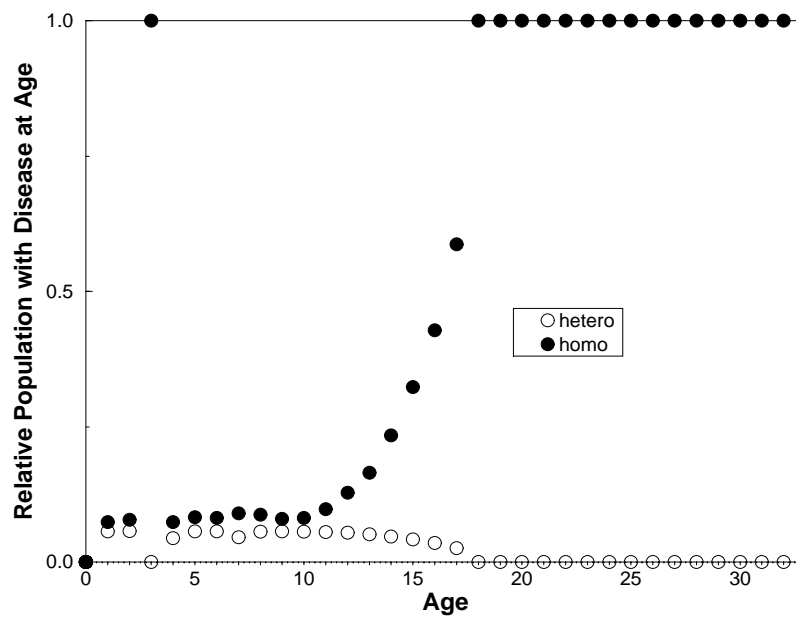
As mentioned before, our results are related to the fixation of relevant bad genes (1 bit) inside the population, after many generations. In Fig. 2 we exhibit, for each position of the genome, the relative number of individuals that have a single bit set to 1 at that position (hetero) or both bits set to 1 at that position (homo). It is important to remember that in both MP and SR cases we consider dominant and recessive positions, while in the AR case all the positions are dominant, that is, 1 bit are always counted. From Fig. 2 we can see that:

- After a given age *greater* than R , all bits are fixed to 1 *in the three reproductive regimes*. This happens due to the dynamics of the model: the well-fitted individuals

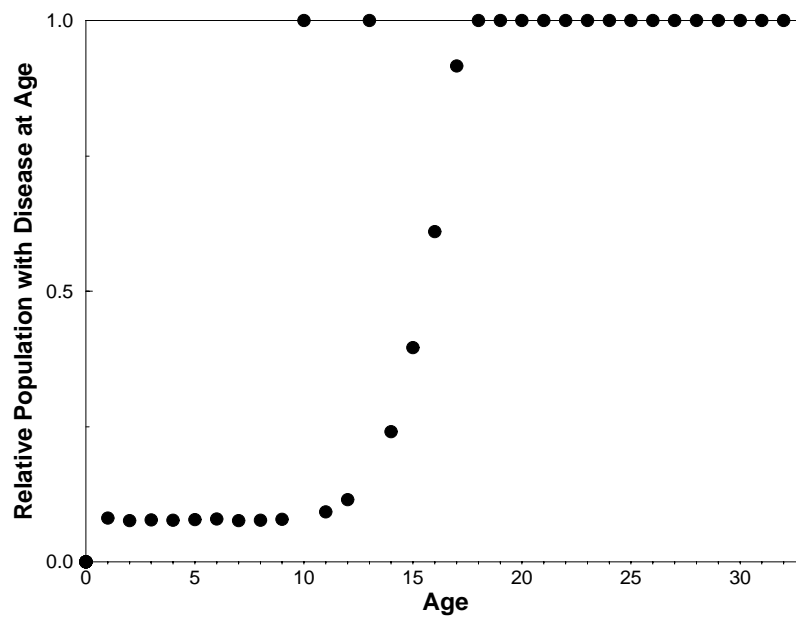


(a)

Fig. 2. Relative number of individuals that have a *single* bit set to 1 (*heterozygotes*, open circles) at that position (age), or *both* bits set to 1 (*homozygotes*, full circles) at that position. Figure 2(a) corresponds to SR, Fig. 2(b) to MP and Fig. 2(c) to AR.



(b)



(c)

Fig. 2. (Continued)

are those with a genome as clean as possible before the minimum reproduction age, since these individuals live enough to generate offspring. Because mutations are unavoidable, *they accumulate at the end part of the genomes*. For this reason these particular fixed genes are often called “irrelevant.” Catastrophic senescence,² which occurs when the whole population breeds only once at the same single age, is closely associated to this fixation at the end part of the genomes. In Fig. 2 these irrelevant bits appear from ages 16 (AR) and 17 (SR and MP) on. Notice that the minimum reproduction age is far below, at age 10 for all cases.

- In SR, the dominant positions are “repulsive”: the number of individuals with those bits set is significantly lower than the average — in the particular simulation shown, bits 4, 7, 17, 26 and 31 were dominant positions. This effect is yet more pronounced for the homozygous individuals: the number of homozygotes with 1 bit *at the dominant positions* is about an order of magnitude smaller than the number of homozygotes with 1 bit in non-dominant positions (Fig. 2(a)). Moreover, apart from the “irrelevant genes,” there is no fixation at all in the SR regime. Such a result is related to the fact that SR has the potential to clean a bad gene in the offspring genome, when a bad fitted parent mates a better fitted one.
- In the MP case *there is fixation in one or two random positions* (for $T = 3$), for the whole population (corresponding to the third bit in Fig. 2(b)). The number of fixed genes varies with the initial random seed. It is important to note that the MP can also clean the offspring genome, although not in a so efficient way as the SR, by the simple fact that SR involves two genomes instead of only one. In the MP if the parent genome is really bad, that of the offspring cannot be much better.
- In AR, *some positions below the minimum reproduction age develop fixation*. In fact, the AR population lives on the edge of death: the number of fixed relevant genes before the reproduction age is *always* equal to $T - 1$, where T is the limit number of allowed diseases. This can be easily understood since in the AR there is no possibility of cleaning a gene.

One aspect that draws immediate attention is that the number of individuals that have a clean gene (relevant part) is greater in the AR case than in the MP case, and the latter one greater than in the SR case — those are the individuals that are *not* represented in Fig. 2. This is related to the fact that in the AR case there is no chance for the offspring to have a cleaner genome than its parent, and only those individuals with a clean genome can have babies with a good probability of survival. In the MP case, there is a chance of cleaning, and in the SR case this chance is the highest of them all. For this reason, the stable population can support a higher number of unclean genomes (heterozygotes) in the SR case without diminishing the survival probability of its offspring. This kind of distribution of genomes is a consequence of the dynamics of the reproductive regime.

It is also important to notice that having a cleaner genome does not mean that AR has a greater survival rate; as mentioned before, it has already been shown and is confirmed by our simulation that the AR survival rate is lower than for the SR,⁴ and that the survival rate of the MP is equivalent to the SR one.⁵ For the AR case, the only possibility of having a stable population in the presence of mutations is to guarantee reproduction by keeping a genome with less than T diseases until reproduction age. However, after reproduction there is a short lifespan since $T - 1$ bits are already fixed before R .

(b) Genetic Diversity and Natural Disaster

In Fig. 3 we show the measured Hamming distance distribution for the three reproductive regimes. The genetic diversity for the AR regime is evidently much smaller than in the other two. Both the SR and MP show Gaussian type distributions with equivalent widths, but the peak of the SR distribution occurs at a longer Hamming distance, corresponding to a higher diversity. We may expect such a result, since the SR is the only regime where there is no fixation of relevant genes at all, although the difference shown is greater than could be expected if this was the only effect present. The presence of heterozygotes in the SR population (Fig. 2(a)) is another reason for a higher diversity, or, the presence of a great number of individuals with a clean genome in the MP case produces a smaller diversity.

In order to verify the importance of this genetic diversity for population survival, we studied the effect of a natural disaster on the time evolution of the populations.

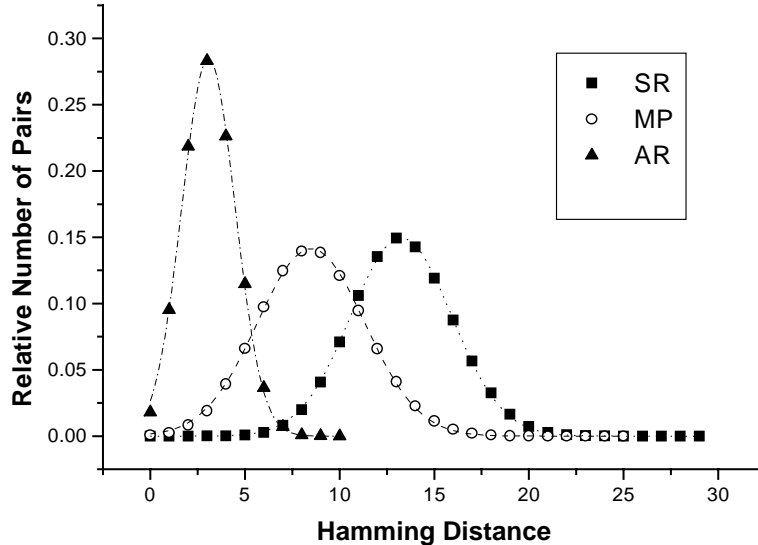


Fig. 3. Hamming distance (genetic diversity) measured after a genetically stable population has been reached, for all cases. Full circles correspond to SR, open circles to MP and triangles to AR.

Table 1. Percentage of the population with a given amount of diseases accumulated at the minimum reproduction age. See text for details.

Diseases at reprod. age	% of the population			
	SR	MP(1)	MP(2)	AR
0	22.56	0	0	0
1	41.36	30.66	0	0
2	27.58	55.82	76.52	79.20
3	7.20	12.68	21.33	20.80
4	1.18	0.84	2.07	0
5	0.12	0.01	0.07	0

To simulate this disaster, we followed the strategy presented by Dasgupta.⁶ After reaching a genetically stable population, one chosen bit was set to 1 (in both bitstrings in the SR and MP cases) in all the genomes. Two different situations can occur: (a) if the chosen bit corresponds to the position of a fixed gene in the AR and MP cases, the time evolution of these populations is not altered; (b) if it corresponds to a non-fixed position in the AR case, meltdown is provoked: the population dies off after a short period — this is a consequence of the fact that the AR population lives on the edge of death, as we mentioned before. In the MP case, the same situation can happen depending on the initial random seed. We noticed that whenever two relevant bits become fixed, mutational meltdown occurs (for $T = 3$). Moreover, we have computed the percentage of the population with a given number of diseases accumulated at the minimum reproduction age R (see Table 1) and noticed that in the MP case, if only one bit is fixed — MP(1) in the table, more than 50% of the population has already accumulated two diseases (for $T = 3$). When there are two bits fixed — MP(2), this number increases to $\approx 75\%$. This means that at least half of the population lives on the edge of death. In the SR case, this number is reduced to $\approx 25\%$, and *meltdown never occurs*. Notice that the numbers in this table are an average performed during the last 250 timesteps of a total of 50 000 steps. For each one of these 250 steps we take a picture of the whole population, and compute the number of diseases accumulated at the reproduction age R , for each individual. In this way even individuals that will not reach reproduction age, and therefore will not generate offspring, are computed. This is the reason why, despite the limit $T = 3$, the table presents also the results for four and five accumulated diseases.

In Fig. 4 we show the results of the natural disaster for the three reproductive regimes.

(c) Cloning

As already explained at the end of Sec. 2, after reaching a stable sexual population, we introduced at one moment in time a given number of female clones into

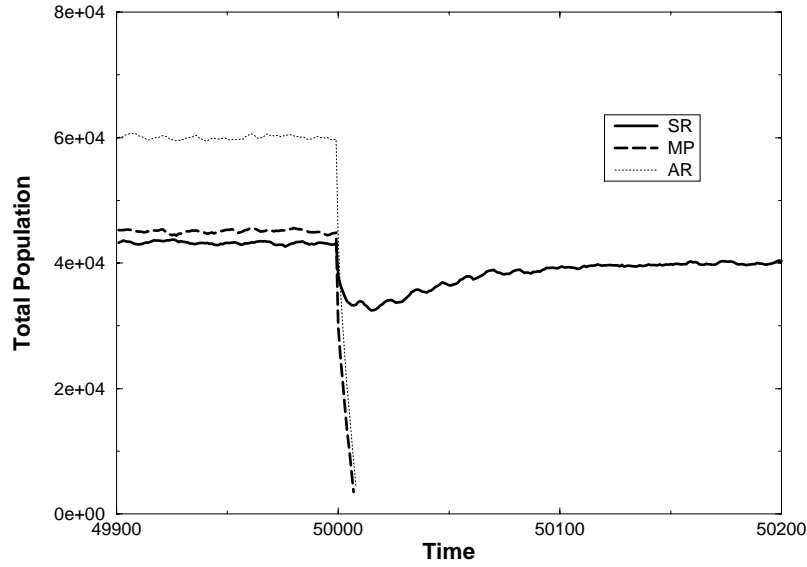


Fig. 4. Evolution of the total population in time, after the genetically steady state has been reached. At step 50 000 the natural disaster is introduced, by setting to 1 a given bit of all the genomes. Notice the extinction of MP and AR populations. Full line corresponds to SR, dashed line to MP and dotted line to AR.

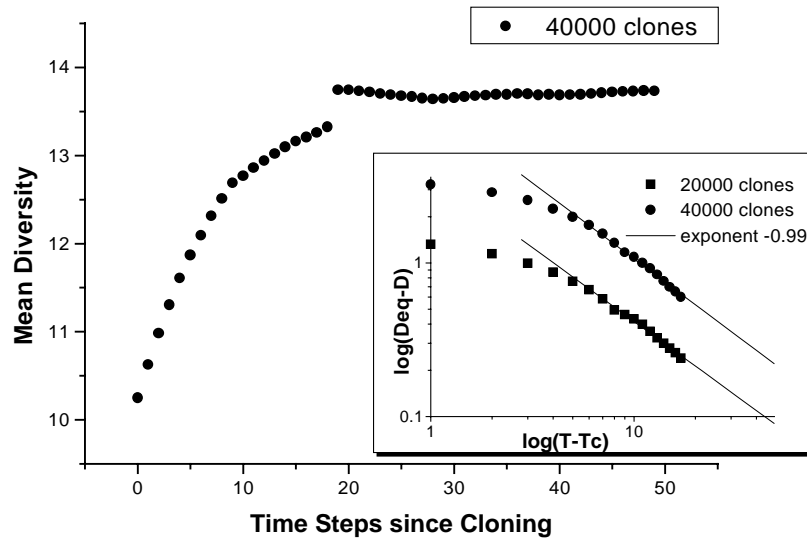


Fig. 5. Evolution of the mean diversity D of an already stable sexual population of 40 000 individuals, after 40 000 females clones are introduced (doubling the population size). The inset shows how the original diversity is restored, for two cases: full circles correspond to the introduction of 40 000 clones, and squares to 20 000 clones. The approach to the original steady state can be fitted by a power law, with an exponent equal to 0.99.

the population. We have measured the genetic diversity immediately before and immediately after cloning and followed its evolution in time. Typical results are shown in Fig. 5. In this figure we plot the mean Hamming distance as a function of the time elapsed since the moment of cloning. As can be observed in the figure, the diversity is restored in a short time, showing that the system loses the genetic memory of cloning very quickly. The approach to the steady state can be fitted by a power law, as shown in the inset of the figure. The exponent of this power law is close to one; it is probably not a coincidence that this same exponent has been previously obtained in the “Eve problem.”⁹ There the authors attributed a family name to each female, which was inherited by the female offspring, and studied the decay of the number of lineages in time. They found that the number of lineages always decay to one as a power law with an exponent also close to one.

The discontinuity observed in Fig. 5 at timestep 17 after cloning is due to the simultaneous death of all the first generation of clones. All these results do not depend on the chosen female genome to be cloned: simulations were made choosing the best fitted, the worst fitted and a random genome, giving always equivalent results. We have also varied the number of clones, from 50% to 100% of the population, without significant differences.

4. Conclusions

We have compared the performance of three types of reproductive regimes: Sexual Reproduction (SR), Meiotic Parthenogenesis (MP) and Asexual Reproduction (AR). We have shown that although the MP and SR present the same survival rates, the genetic diversity of SR is larger, decreasing the possibility of extinction when a natural disaster occurs, while AR and MP did not. We have related this genetic diversity to the fixation of bad relevant genes inside the population genomes. We have shown that simple asexual reproduction turns out to be the worst reproductive regime, with the smallest survival rate and genetic diversity. In this case the fixation of bad relevant genes drives the population to live on the edge of death. The SR regime does not present fixation of relevant genes at all, and has the highest genetic diversity. The MP regime is in between the other two.

We have also studied how the genetic diversity of an already genetically stable sexual population is restored after a given amount of female clones has been introduced. The restoration of the steady state follows a power law in time, with an exponent close to one.

Acknowledgments

We acknowledge P. M. C. de Oliveira for useful discussions and the Brazilian agencies CNPq, CAPES, FAPERJ and FINEP for financial support.

References

1. T. J. P. Penna, *J. Stat. Phys.* **78**, 1629 (1995).
2. T. J. P. Penna, S. Moss de Oliveira, and D. Stauffer, *Phys. Rev.* **E52**, 3309 (1995);
T. J. P. Penna and S. Moss de Oliveira, *J. Physique* **I5**, 1697 (1995).
3. S. Moss de Oliveira, P. M. C. de Oliveira, and D. Stauffer, *Sex, Money, War and Computers* (Springer, Singapore, submitted 1997).
4. D. Stauffer, P. M. C. de Oliveira, S. Moss de Oliveira, and R. M. Zorzenon dos Santos, *Physica* **A231**, 504 (1996).
5. A. T. Bernardes, *J. Stat. Phys.* **86**, 431 (1997).
6. S. Dasgupta, *Int. J. Mod. Phys.* **C8**, 605 (1997).
7. P. Pamilo, M. Nei, and W. H. Li, *Genet. Res., Camb.* **49**, 135 (1987).
8. S. C. Stearns, *The Evolution of Sex and its Consequences* (Birkhauser, Basel, 1987).
9. P. M. C. de Oliveira, S. Moss de Oliveira, and D. Stauffer, *Theory Biosci.* **116**, 3 (1997).