

## GAMETE RECOGNITION AND COMPLEMENTARY HAPLOTYPES IN SEXUAL PENNA AGEING MODEL

S. CEBRAT<sup>†</sup> and D. STAUFFER<sup>\*</sup>

*Department of Genomics, Wrocław University  
ul. Przybyszewskiego 63/77, 51-148 Wrocław, Poland*

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

*<sup>†</sup>cebrat@smorfland.uni.wroc.pl*

*\*stauffer@thp.Uni-Koeln.DE*

Received 25 September 2007

Accepted 8 October 2007

In simulations of sexual reproduction with diploid individuals, we introduce female haploid gametes that recognize one specific allele of the genomes as a marker of the male haploid gametes. They fuse to zygotes preferably with male gametes having a different marker than their own. This gamete recognition enhances the advantage of complementary bit-strings in the simulated diploid individuals, at low recombination rates. Thus with rare recombinations the bit-strings evolve to be complementary; with recombination rates above approximately 0.1 they instead evolve under Darwinian purification selection, with few bits mutated.

*Keywords:* Monte Carlo simulation; sexual reproduction; haplotype selection.

Darwinian evolution is usually thought to come from a counterplay of two opposing trends:

- (i) survival of the fittest, which purifies the genome from bad mutations, and
- (ii) random accidents that produce new bad mutations.

The Penna ageing model of mutation accumulation, as reviewed in detail elsewhere,<sup>1</sup> is one such example. Bad mutations are stored in strings of bits, they reduce the survival chances of the adults, and new bad mutations may happen at birth. Then a realistic mortality is obtained, increasing exponentially with adult age, for both asexual and sexual reproduction.

Recently, however, several simulations of sexual reproduction<sup>2–5</sup> found another strategy: the two bit-strings of sexual (“diploid”) animals may become complementary. Wherever one bit-string has a zero, the other has a one, and vice versa. If all mutations are recessive, and no genetic loci are dominant, then this combination of bit-strings is as favorable as if both bit-strings would contain only zeros. Most of the bit-strings in the equilibrium population are then divided into two groups, A and

B, such that each diploid individual has one A and one B bit-string. After sexual reproduction mixes the paternal and maternal genomes, only those zygotes with one A and one B bit-string can survive for long. This ideal picture is made fuzzy due to new mutations (at birth), some dominant loci, and recombination (crossover) of the two bit-strings. Thus the emergence of complementary bit-strings from initially mutation-free genomes happens at low but not at high recombination rates  $r$ .<sup>2-6</sup> At high  $r$  above approximately 0.1, purification instead dominates, which means most bits are zero.

In the present work we continue Ref. 2 and introduce the selection of good male gametes in sexual reproduction. It was claimed<sup>4,7</sup> that recognition of the major histocompatibility complex (MHC) guides females in the selection of males. This, however, seems to be unrealistic; we know that the type of car, the size of the residence, and the bank accounts are crucial selection criteria but were ignored in all simulations. Instead we deal with recognition and selection on the level of gametes having only one bit-string (“haploid”). We assume that the haploid female gamete (“ovum”) recognises a “marker”, i.e. a part of the genome in the haploid male gamete (“sperm”), before fusion into a zygote takes place. Thus a B ovum prefers to select an A sperm, and vice versa, somewhat similar to antibody–antigen reactions in immunology. This marker corresponds to the “driver” in Ref. 2.

More precisely, the first of the  $L$  bits in a bit-string determines gamete preference. Before fusion, with probability  $s$ , a female gamete tries to fuse with a male gamete having the opposite first bit. If the ovum has zero as first bit, it tries to find a male gamete that has one as the first bit, and vice versa. If the first attempt is unsuccessful, it is repeated until a gamete with an opposite first bit is found. All repetitions involve the male gametes of the same male partner, but these gametes undergo the usual mutations before they are tested.

In our standard case, the carrying capacity  $N_{\max}$  for the Verhulst factor (applied to births only, not to adults) was 2000; the bit-string had a length  $L = 64$ ; at each iteration  $128/L$  births happened per adult female; the minimum reproduction age was  $5L/8$ ;  $T = 1$  active mutations killed the individual; at birth one random irreversible mutation (from good = 0 to bad = 1) was tried in each gamete; if the selected bit was already one, nothing changed. Survival was tested for  $t = 10^5$  iterations (in one iteration all individuals aged by one time unit and could reproduce, if old enough). Gamete recognition and selection is switched on after one quarter of the simulation time  $t$ . For the standard aspects of the Penna model we refer to Ref. 1 and numerous articles in this journal, like Ref. 3. We now look at the dependence on the crossover (recombination) probability  $r$  between 0 and 1, and systematically vary various input parameters. We check how many of 100 samples survive to the end of the simulation; this number is then the survival probability in percent.

To measure the degree of complementarity we looked at the Hamming distance, which is the number of positions on which the bits of the male and the female gamete differ in a position-by-position comparison at the moment of fusion into a zygote.

Only the  $5L/8$  (usually 40) bits for young age were compared since the bits beyond the minimum age of reproduction are mostly mutated to one after equilibrium has been reached. Thus the Hamming distance varies from 0 (full agreement) to 40 (full complementarity). We sum it up over the second half of the simulation, i.e., over 50 000 iterations.

Figure 1 shows our standard case (plus signs) as well as one with a higher population and another one with a higher birth rate. For the higher population, Fig. 2 gives the resulting distribution of the Hamming distance. As in Ref. 2, at

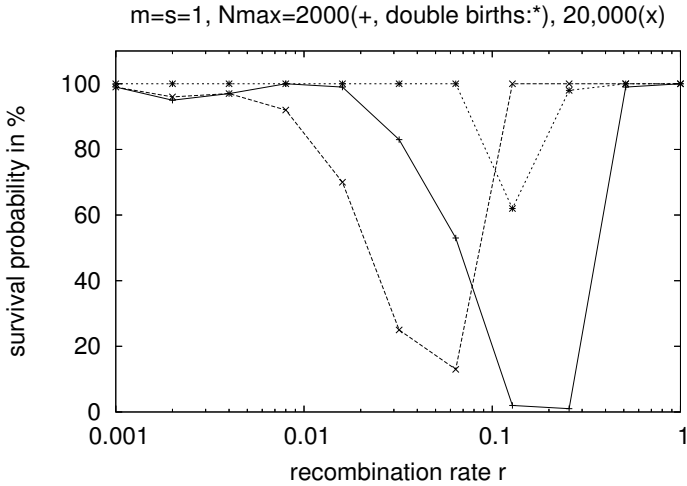


Fig. 1. Number, from a hundred runs of samples, where the population survived for at least  $t = 10^5$  iterations. The + refers to our standard parameters; for × the populations were ten times higher, and for the stars instead the birth rate was doubled from 2 to 4.

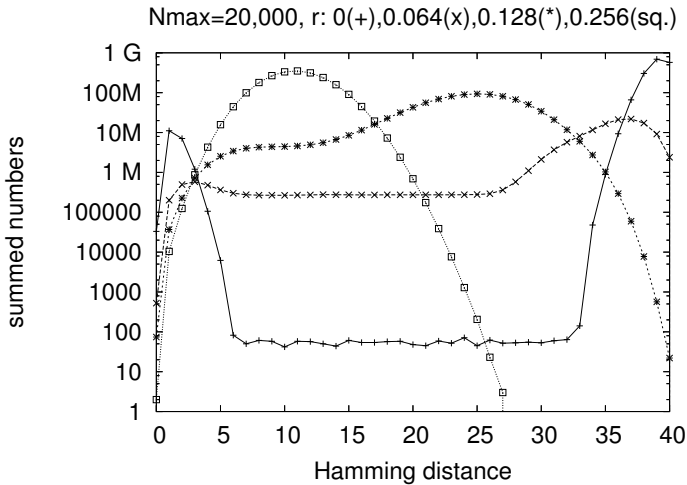


Fig. 2. Distribution of Hamming distances for one of the simulations of Fig. 1.

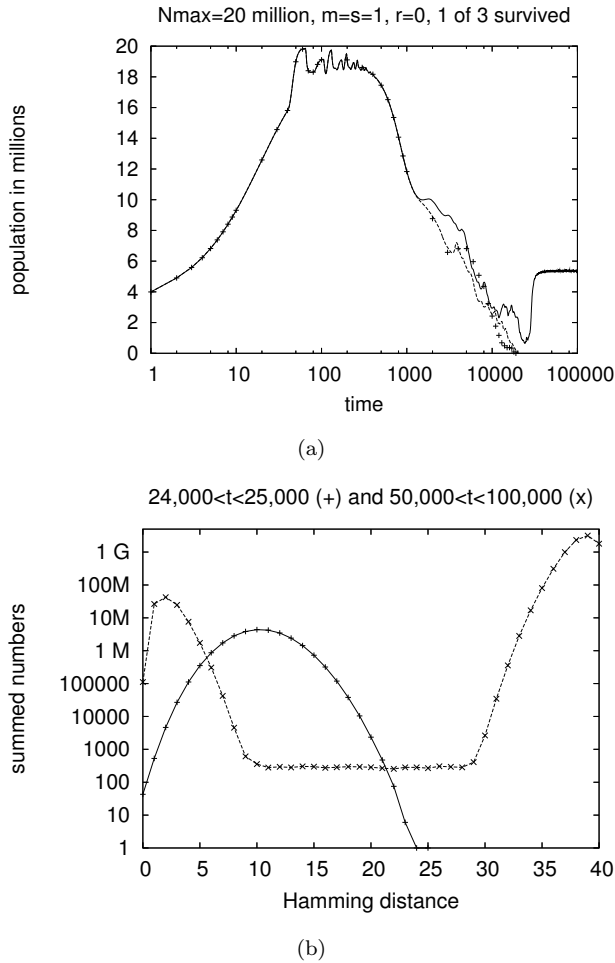


Fig. 3. (a) Time variation of the population, and (b) of the Hamming distance of a large population.

intermediate  $r$  survival is difficult. For smaller  $r$ , complementarity dominates with Hamming distances near 40 (with about half the bits set to zero = correct allele), while for higher  $r$  the two bit-strings are much more similar and most bits are zero. Higher populations had difficulties surviving long enough, but after two failed attempts we succeeded in one with about 5 million individuals, for  $N_{\max} = 20$  million (see Fig. 3). The Hamming distances are small (not shown) and the population low (Fig. 3) shortly before gamete recognition is switched on, while the population rapidly increases when gamete recognition is switched on.

When we reduce the gamete recognition probability  $s$ , survival at low  $r$  becomes more difficult (see Fig. 4); the corresponding curves are more equidistant on the left part if, instead of the number of the surviving samples, we plot the average population (not shown). Increasing the number  $d$  of dominant loci from the standard

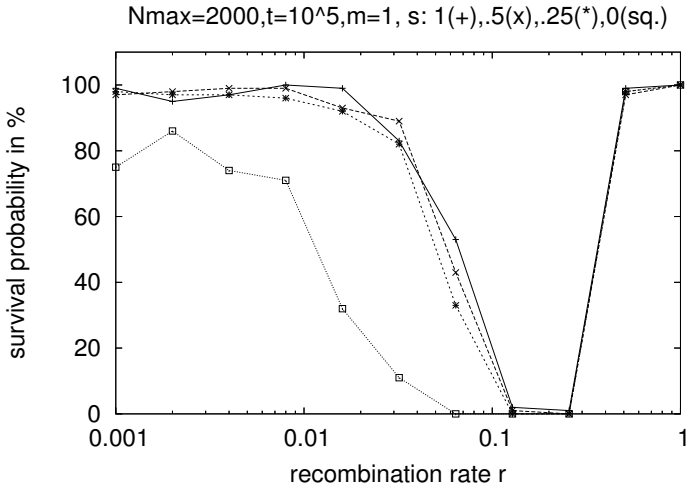


Fig. 4. Dependence of the gamete recognition probability  $s$ .

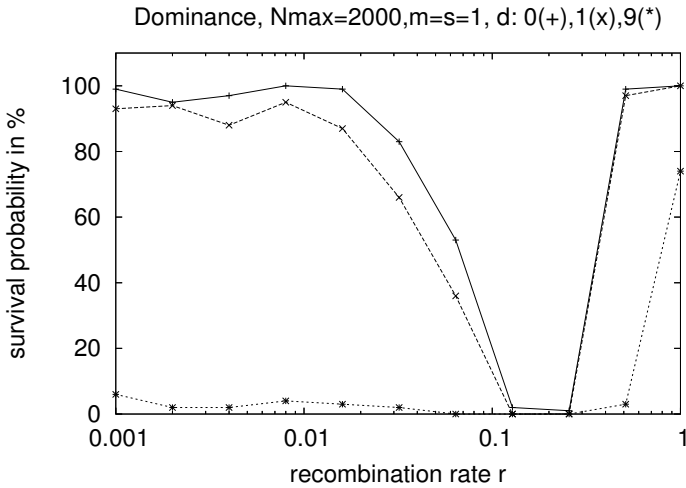


Fig. 5. Dependence on the number  $d$  of dominant bits.

case (where  $d = 0$ ) has little effect when  $d = 1$ , but is disastrous when  $d = 9$  at low  $r$  (see Fig. 5). Surprisingly, a slight reduction of the mutation probability  $m$  from the standard (where  $m = 1$ ) increases the survival at intermediate  $r$  drastically (see Fig. 6).

The following results are not shown as figures. When we reduce the length of the bit-strings (from 64 to 32 and 16), or the minimum reproduction age (from 40 to 24), the minimum of the survival probability becomes less pronounced and shifts to smaller  $r$ . When the observation time is increased from  $10^5$  to  $10^6$ , the survival chances go down drastically except for  $r = 1$ ; in principle, finite populations (with

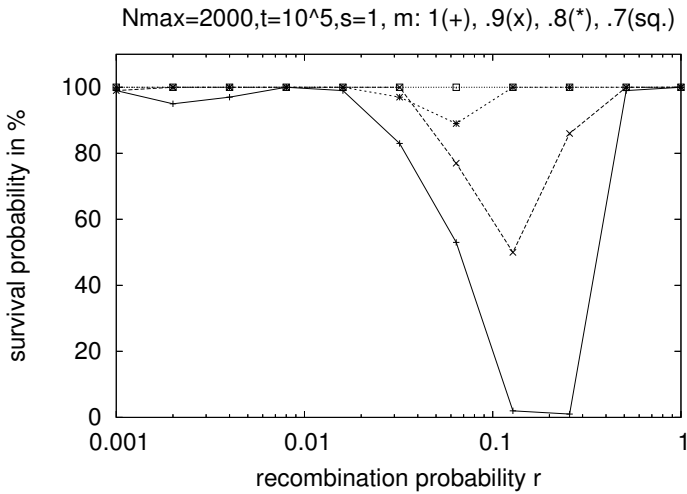


Fig. 6. Dependence on the mutation probability  $m$  per iteration.

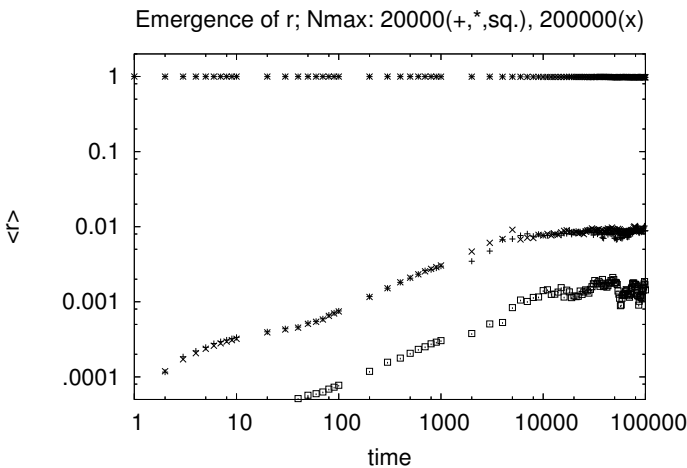


Fig. 7. Emergence of a self-organizing average recombination rate. Results for  $N_{\max} = 2$  million were similar (not shown).

$10^3$  individuals like ours) always vanish if simulated over sufficiently long times. The distinction between small and large  $r$  in the distribution of Hamming is always observed. If the lethal threshold  $T$  for the number of mutations is increased from 1 to 2 and 3, all samples survive, and the complementarity is less pronounced. Thus our results remain only qualitatively the same if we change parameters.

Finally we looked at the emergence of an optimal recombination rate  $r$  in the standard way.<sup>8</sup> Each individual has its own  $r$ , initially 0 for all, or 1 for all. Then at birth sons inherit their  $r$  from the father, and daughters from the mother, apart from a random change  $\pm 0.001$ . Figure 7 shows the average (of all females) first

to increase from zero up to about 0.01 (middle curve); if instead we start with  $r = 1$  for all, the average remains above 0.95 (upper data). [If we reduce the change from  $\pm 0.001$  to  $\pm 0.0001$ , the final average is also reduced by this factor ten (lower curve)]. Thus, the recombination rate  $r$  “wants” to stay at either zero or one, but mutations hinder it from staying at these ideal values.

In summary, we confirmed the separation of two evolutionary strategies,<sup>2</sup> bit-string complementarity at a low recombination rate and purification selection (few mutations) at a high recombination rate, for different parameters. In particular, our newly introduced gamete selection enhances the usefulness of this complementarity (Fig. 4).

### Acknowledgments

We thank COST P10 for supporting the visit of D. Stauffer.

### References

1. S. M. de Oliveira, P. M. C. de Oliveira and D. Stauffer, *Evolution, Money, War and Computers* (Teubner, Stuttgart–Leipzig, 1999); D. Stauffer *et al.*, *Biology, Sociology, Geology by Computational Physicists* (Elsevier, Amsterdam, 2006).
2. M. Zawiarta *et al.*, *Theor. Biosci.* **125**, 124 (2007).
3. K. Bońkowska *et al.*, *Int. J. Mod. Phys. C* **18**(8), 1329 (2007).
4. A. Pękałski, *Int. J. Mod. Phys. C* **18**(10), 1919 (2007).
5. P. M. C. de Oliveira *et al.*, Does sex induce a phase transition? (Sept. 2007), preprint.
6. W. Waga *et al.*, *Theor. Biosci.* **126**, 53 (2007).
7. S. Jacob *et al.*, *Nat Genet.* **30**, 175 (2002).
8. N. Ito, *Int. J. Mod. Phys. C* **7**, 107 (1996).