The population genetics of parasite-host systems and the preservation of genetic diversity

K. James Soda Major Advisor: Dennis E. Slice, PhD

Department of Scientific Computing, Florida State University

Introduction: Ever since it was first proposed by Haldane (1949), there have been theories asserting that parasite-host interactions allow genetic diversity to persist. Clarke (1976) discussed one mechanism by which this may occur. Consider a host population that has two alleles, A and B, which determine their susceptibility to a parasite. In turn, the parasite population has two alleles, A' and B'. Parasites with allele A' are only able to infect hosts with allele A, and parasites with allele B' are only able to infect hosts with allele B. If allele A is very common in the host population, selection will favor A' in the parasite population. As a result, the frequency of A' will increase. However, as A' becomes more common, selection will favor allele B in the host. Therefore, the allele frequency for A will decrease. Then, as A becomes less common, A' will become less common as well. As a result, the parasites and hosts will "loop" through which allele is the most common, and neither allele will become fixed (Clarke 1976). If we were to plot the allele frequencies of A and A' over time, this looping process should appear similar to a sinusoidal curve. The goal of this study was to generate a simulation mimicking the sort of parasite-host system described above, assess whether this simulation provided any evidence for the looping mechanism described in Clarke (1976), and determine whether genetic variability can be maintained for a substantial amount of time in this system.

Methods

Basic Model Structure

The simulation software was written in Python ver 2.7 (www.python.org), and all graphics were generated using the matplotlib library (Hunter 2007). The hosts were assumed to be asexual, haploid organisms with discrete generations. Every host was assigned one of two alleles with equal probability of both. In addition, every host had a set probability of becoming infected with a parasite. If the host was infected, the parasite was assigned an allele of its own. For each host allele there was a corresponding parasite allele, each occurring at the same frequency. As a result, each host could be classified into one of three categories: uninfected, infected with a parasite whose allele matches its own (henceforth referred to as a matching host), and infected with a parasite whose allele does not match its own (henceforth referred to as a non-matching host).

At the end of the first generation the hosts contribute offspring to an infinitely large pool. The proportion of offspring from any given individual depends on both the category to which the individual belongs and to which group other members of the population belong. A new generation of hosts was then created by randomly sampling from this infinite pool until the new generation had the same population size as the previous generation.

In addition, the parasite allele frequency was updated at the end of the generation. The new allele frequencies were the weighted proportion of parasites with each allele in the current generation. Once a new parasite allele frequency was calculated, it was used to infect the next generation of hosts. This process was repeated over several generations.

Simulations

The simulation assessed whether the previously described looping pattern arises under the conditions of the model. Populations of 1000 host individuals were followed for either 1000 generations or until both the host and the parasite became fixed for one allele. At the end of each run, the simulation plots the frequency of both host and parasite allele A against time. I then qualitatively assessed whether these plots showed evidence for the looping mechanism based on the presence of a sinusoidal pattern. To ensure that the patterns observed in the simulation results were indicative of a large scale trend and not simply the idiosyncrasies of a single run, I ran the simulation ten times and compared the resulting plots. I used two infection rates. At the low infection rate, each host had a 10% probability of becoming infected while at the high infection rate each host had a 90% chance of becoming infected. Table 1 reports the remaining parameters used in the model.

Results: When the infection rate was low, the majority of simulation outputs resembled a random walk pattern and showed no clear evidence for the looping mechanism suggested in Clarke (1976) (Figure 1). However, occasionally the simulation generated data resembling Figure 2. Although the curves still portray a great deal of noise, there are some regions that resemble a cyclic pattern. It is feasible that the looping mechanism is present within these regions.

In contrast, when the infection rate is high the simulation output tends to look similar to Figure 3. Here the sinusoidal pattern associated with the looping mechanism is clearly present. However, the amplitude of the curves becomes gradually greater until one allele reaches fixation.

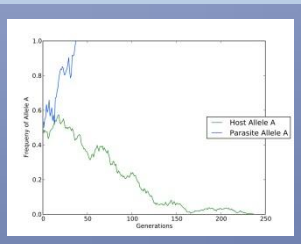
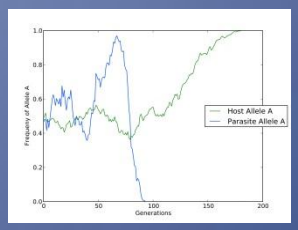


Figure 1. A typical result when infection rates are low.

Figure 2. A result from when the infection rate is low where there may have a weak sinusoidal pattern



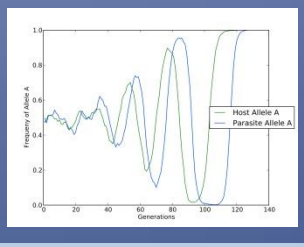


Figure 3. A typical result from Simulation2 when infection rate is high.

Discussion: Based on how allele frequencies change through time, the looping mechanism described in Clarke (1976) is theoretically possible. However, this mechanism does not appear to have a notable impact on the population genetics of the system unless the infection rate is relatively high. Even in instances where the infection rate is high, though, the looping mechanism does not maintain genetic variability over a sustained period of time. Therefore, at least under the assumptions of this simulation, parasite-host interactions should not be expected to maintain variability. These results agree with Yu (1972), which found that genetic variability cannot be maintained via this mechanism if the either the parasite or the host is haploid.

References

Clarke, B. 1976. The ecological genetics of host-parasite relationships. In *Genetic Aspects of Host-Parasite Relationships*. AER, Taylor and R Muller (eds). pg. 87-103. Blackwell Scientific Publishing, Oxford

Haldane, JBS. 1949. Disease and evolution. Supplement to La Ricerca Scientifica 19: 68-76 (retrieved from http://www.ff.unair.ac.id/ebooks/Malaria%20Gen etic%20and%20Evolutionary%20Aspects.%20Dro namraju,%202006.pdf#page=183 11 Dec 2011)

Hunter, JD. 2007. Matplotlib: A 2D graphics environment. Computing in Science and Engineering 9: 90-95.

Yu, P. 1972. Some host parasite genetic interaction models. Theoretical Population Biology 3: 347-357.

Table 1. Parameter values used in the simulation.

Parameter	Value
Initial A Allele Frequency (Host)	0.5
Initial A Allele Frequency (Parasite)	0.5
Ratio of Uninfected Reproductive Output: Matching Host Reproductive Output: Non-matching Host Reproductive Output	4:2:1
Weight Given to Parasites in Non-matching Hosts (Both A and B)	0.5