

## PopGen Fishbowl: A Free Online Simulation Model of Microevolutionary Processes

THOMAS C. JONES,  
THOMAS F. LAUGHLIN



### ABSTRACT

Natural selection and other components of evolutionary theory are known to be particularly challenging concepts for students to understand. To help illustrate these concepts, we developed a simulation model of microevolutionary processes. The model features all the components of Hardy-Weinberg theory, with population size, selection, gene flow, nonrandom mating, and mutation all being demonstrated in the simulations. By using this freely available computer model, students can develop and test hypotheses with replicated virtual experiments. Because the model is an agent-based simulation, there is biologically realistic variability in the results. Students using the model see results both numerically and graphically, and these are reinforced by an animation of the virtual fish in the simulated experiment.

**Key Words:** Evolution simulation; natural selection; Hardy-Weinberg equilibrium; gene flow.

Computer models have several features that help in teaching the dynamics of basic concepts in population genetics. Although there is no substitute for a firm understanding of the basic underlying quantitative features of the Hardy-Weinberg equilibrium and the resulting effects of violations of its assumptions, it can be very difficult to examine the long-term dynamics of population genetic processes with only a pencil-and-paper approach. It is also practically impossible to experimentally examine such biological processes in the typical college-level course with the usual limitations of time and resources. On the other hand, it is very important for students to understand the relationship between the process of microevolution at the population level and the changes in gene pools caused by violations of the Hardy-Weinberg assumptions. Computer models that demonstrate the long-term dynamics of population genetic processes can be very valuable to instructors in this regard.

Perhaps more than any other discipline within biology, evolution is conceptual in nature and cannot be mastered by rote. Students must have a strong working understanding of the concepts to be able to apply them. For example, a student might memorize that gene flow and population size affect the possibility of genetic drift in a population, but unless students understand both concepts and how they interact, they will not be able to analyze whether or not significant changes in allele frequencies are to be expected in a given scenario. In

*Evolution is identified  
as one of the most  
misunderstood and  
difficult concepts to teach  
in biology.*

fact, evolution is identified as one of the most misunderstood and difficult concepts to teach in biology (Bishop & Anderson, 1990; Ferrari & Chi, 1998).

Citing examples and case studies in lecture is useful when teaching evolutionary concepts, but this does not require students to engage directly with the material. By contrast, conducting experiments in laboratories challenges students to formulate hypotheses and make predictions based on theory, which facilitates conceptual learning. The problem, unfortunately, is that realistic evolutionary experiments typically take much longer than a semester or academic year. As a result, it has been difficult to develop practical educational analogues for most classic experiments in lab and classroom settings that foster an understanding of theory through empirical demonstrations.

Although they are by no means replacements for hands-on laboratory or field experience, computer simulations provide a useful tool for illustrating and conveying complex concepts in evolution (Soderberg & Price, 2003; Latham & Scully, 2008; Perry et al., 2008). In the most realistic simulations, the interactions among individual organisms are modeled with elements of probability and randomness. The interactions of simulated individuals (agents) generate results that are affected by variables in a probabilistic, rather than deterministic, way. In other words, agent-based simulation models create a virtual world in which experiments can be conducted. Moreover, because of the built-in stochasticity, such simulations generate data that are much more biologically realistic and interesting than deterministic models that generate results from mathematical formulae.

In the virtual world, students can quickly conduct experiments that would take years or even decades in the real world. By working with simulations, students are obliged to interact with the concepts being illustrated. They cannot be passive. Students learn the effects of different variables by tweaking them and observing the results. Furthermore, students will see that biological results are not as neat and tidy as the graphs typically presented in textbooks. Rather, real biological results are messy, and trends in the data are revealed by statistical analyses.

The freeware models available now are limited in their usefulness for describing microevolution in some significant aspects. One of the

typical limitations is that most of these models make deterministic predictions of the effects of most population genetics forces, such that outcomes of a particular combination of parameters will not vary over any number of model runs (e.g., the models Populus and Visual PopGen). Population genetics computer simulations also typically require installation of the software on each individual computer in use, which can use up lots of classroom and student study time just to have the simulation operate properly (e.g., Populus, Visual PopGen, and Critters!). PopGen Fishbowl is a NetLogo model (Wilensky, 1999) that runs from the Internet inside a browser window, without the need to download and install the software on the user's computer. Our model is freely available at <http://faculty.etsu.edu/jonestc/Virtualecology.htm>. The only requirement is that the Web browser be updated with the common plug-in Java 1.4.1.

PopGen Fishbowl gives students the ability to simulate the effects of deviations from all the Hardy-Weinberg assumptions either independently or in any particular combination. The model also allows students to simulate other important population parameters such as sex ratio, fecundity, survivorship, and carrying capacity. The model produces results that are numerical, graphic, and visual (animated fish with variable phenotypes) to aid students in understanding the dynamics of population genetic processes.

The advantage of this model over many others is that students get to observe dynamic changes within the population. For example, allele frequency changes are shown not only as a value or graph but also as individual phenotypes in the simulation's world view (Figure 1). Being agent-based means that in addition to producing more biologically realistic variability, the model also provides an opportunity for students to perform statistical analyses of the simulation results.

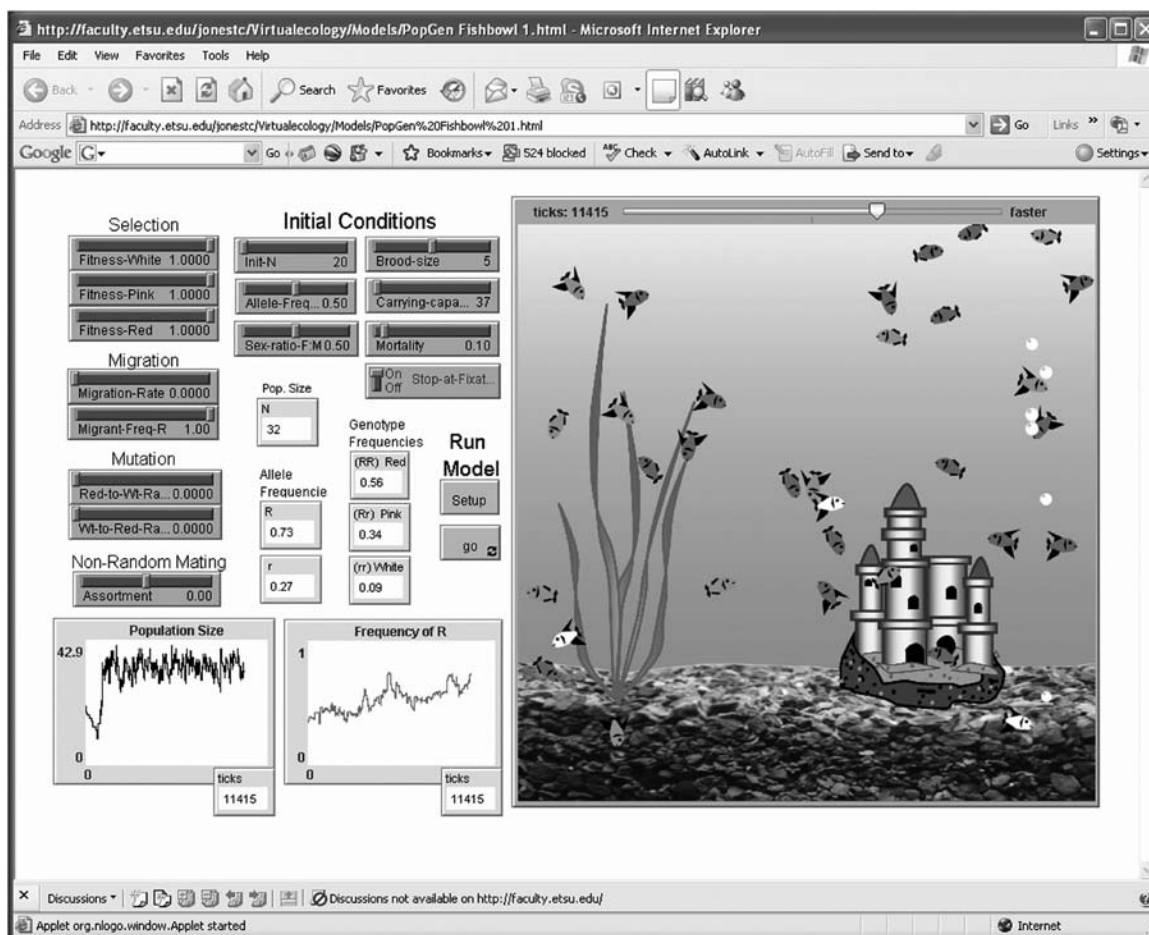
## ○ Model Details

The agents of the model are fish in a small "aquarium" population. The frequency of a single gene with two alleles affecting body color is tracked over time in the simulated aquarium population. The two alleles (red and white) are incompletely dominant, and the heterozygotes are pink. When fish of opposite sex come together, they may breed and produce offspring that follow patterns of Mendelian inheritance in their body color. Overlapping generations are simulated in the model because reproduction can occur whenever fish meet.

Virtual experiments can be done by varying the population's initial conditions of allele frequencies, population size, carrying capacity, sex ratio, mortality, and brood size (Table 1). In addition to population size, the Hardy-Weinberg parameters of relative fitness, migration (gene flow), mutation, and assortative mating can be set at the beginning of each simulation run (Table 2). By varying these parameters, students can explore the effects of violating Hardy-Weinberg assumptions regarding population size, selection, mutation, migration, and nonrandom mating. Some basic population biology such as logistic growth and the Allee effect can also be explored by manipulating different parameters of the model.

A run of the simulation can be either started or stopped by selecting the "go" button, and new parameter values are implemented by selecting the "setup" button. Initial conditions and experimental variables can be changed by the user with a set of sliders in the user interface (Figure 1). Both the initial population size and the carrying capacity can vary from 0 to 1000. The brood size can range from 0 to 10. The sex ratio is set as the probability that any particular fish will be female when born and can vary from 0 to 1. Each simulated fish in the population has a probability of dying within 100 time intervals ("ticks"). The mortality probability

Figure 1. PopGen Fishbowl user interface running in web browser.



**Table 1. Population parameters that can be varied in the PopGen Fishbowl simulation.**

|                                |  |
|--------------------------------|--|
| <b>Initial Population Size</b> | Can range from 0 to 1000   |
| <b>Allele Frequency</b>        | Frequency of R (range: 0 to 1), with $r = 1 - R$                               |
| <b>Sex Ratio</b>               | Probability that any individual is female (range: 0 to 1)                      |
| <b>Brood Size</b>              | Can range from 0 to 10   |
| <b>Carrying Capacity</b>       | Can range from 0 to 1000   |
| <b>Mortality</b>               | Probability of each individual dying within 100 time intervals (range: 0 to 1) |

**Table 2. Experimental variables that can be manipulated in the PopGen Fishbowl simulation.**

|  |   |
|--|---|
| <b>Selection</b><br>Fitness – white<br>Fitness – pink<br>Fitness – red | Relative fitnesses of white, pink, and red phenotypes (range: 0 to 1)   |
| <b>Gene Flow</b><br>Rate<br>Frequency of R                             | Rate of migrants carrying genes into the population (range: 0 to 0.1) and the frequency of allele R in the migrating population (range: 0 to 1) |
| <b>Mutation</b><br>Red to white<br>White to Red                        | Rates of mutations and back mutations in juvenile fish (range: 0 to 0.01)   |
| <b>Nonrandom Mating</b>  | Likelihood of positive or negative assortative mating (range: –1 to 1)  |

selected by the user is a maximum that adjusts proportionally to the population size divided by carrying capacity. Mortality reaches the maximum value when population size is equal to the carrying capacity. The two alleles in the population are R and r, with the genotypes RR (= red body color), Rr (= pink), and rr (= white). Each simulation run is initiated by specifying the proportion of R alleles in the population using a slider on the simulation's user interface. The effects of selection are simulated by choosing a relative fitness (range: 0–1), which is imposed on juveniles prior to reproduction during the simulation. Juveniles become adults after 10 ticks. Note that the simulation of selection against a dominant or recessive allele in the model can be accomplished by making the relative fitness of the heterozygote Rr equal to that of one of the homozygotes. The rate of migration into the population is implemented by a probability (range: 0–0.1) that a fish from some other population replaces any one fish in the simulated population. The probability that an immigrating fish carries either the R or r allele is set by adjusting the frequency of R in the immigrant population (range: 0–1). Mutations (R to r) and back mutations (r to R) are simulated as separate probabilities that each allele in a juvenile will mutate (range:  $10^{-4}$  to  $10^{-2}$ ). Nonrandom mating is simulated by using a slider on the interface to adjust the probability of positive assortative mating (value > 0) or negative assortative mating (value < 0), such that individual fish have a greater probability of mating with individuals of the same or different genotype than expected by chance (value = 0). Using a toggle in the interface, the user can have the simulation either stop or continue when one allele becomes fixed (R = 1 or 0). The simulation stops running if population size becomes zero.

Output from the model is both numerical and graphical (Figure 1 and Table 3). Population size, allele frequencies, and genotype frequencies are reported numerically. Continuously updated graphs show both population size and the frequency of the R allele over time. The model also produces a continuously updated image that shows the sex and body-color phenotype of each fish in the aquarium population.

The rate at which the model runs can be adjusted with a slider located at the top of the world view in the user interface. “Explosions” of fish fry

from individual breeding events can be observed at very slow model running speeds. At very high speeds, the simulation world view will regularly display snapshots of the population, and changes are more easily monitored in the graphic output. This feature is useful in speeding up virtual experiments. The simulation can also be paused at any time during a session to produce a snapshot of population data. If the probability of mortality is set to 0, the population will then cease to evolve once the population size reaches the carrying capacity because there will be no further reproduction. Thus, higher mortality rates and larger brood sizes will increase the rate of evolution in the population. Various other dynamics related to demography can also be observed in the simulated population by adjusting the non-Hardy-Weinberg parameters that are included in our model. For example, if the initial population density is set very low, the population may become extinct because individuals fail to find a suitable mate (Allee effect). If the user specifies a greatly unbalanced sex ratio, this can cause the population to become extinct because of the increased likelihood that one of the sexes will become absent in the population at some time in the future.

## ○ Using Fishbowl in the Classroom

With the PopGen Fishbowl model, students can use an agent-based model to investigate all the effects of violations of assumptions in the Hardy-Weinberg equilibrium in a biologically realistic simulation. The effects of nonrandom mating, different population sizes, or changes in the magnitude of selection, mutation, or migration can be evaluated independently or in combination with one or more of the other forces. Because this is an agent-based simulation model rather than a deterministic model, there is variability in the effects of parameter settings on the outcome of any particular run of the model. Therefore, students can observe some of the variance in observed outcomes that are expected in real populations.

We use PopGen Fishbowl to allow students to first make, and then test, hypotheses about the effects of varying a single Hardy-Weinberg parameter in a simulated population. Students make predictions regarding allele and genotype frequencies based on their understanding of the Hardy-Weinberg

**Table 3. Output produced by PopGen Fishbowl simulation.**

|                             |  |
|-----------------------------|--|
| <b>Population Size</b>      | A numerical output and a continuously updated graph of population size over the total time course of the simulation.   |
| <b>Allele Frequencies</b>   | A numerical output of proportions for two alleles (R and r) and a continuously updated graph of the proportion of one allele (R) over the total time course of the simulation. |
| <b>Genotype Frequencies</b> | Continuously updated numerical output of three genotype frequencies RR, Rr, and rr.  |
| <b>Simulated Population</b> | Animated representation of the complete population showing each individual with phenotypes for body color and sex.   |

equilibrium theory. The simulation model generates data that students then analyze statistically. Testing these predictions with the simulation-model-based experiment reinforces the basic principles covered during the introduction of the material to students during lecture. Students are generally free to make and test their own hypotheses and carry out repeated simulations with the same parameter settings to gain experience with the variability expected in natural populations. Replicated model runs allow students to estimate means and variances of population measures that can then be used for testing their own predictions. For example, a student could run two sets of simulations to test for the effects of gene flow rates on allele frequencies. In one set of simulations the student could set a given migration rate, while in a second set of simulations migration could be set to zero. Using means and variances of allele frequencies from the resulting simulations, the student could test for significance of the effect of the hypothetical migration rate on the evolution of allele frequencies in the simulated population. For a more robust experiment, an entire class could cooperate to produce simulation data for testing an agreed-upon hypothesis.

As a second step, students can make predictions of the effects of violating more than one Hardy-Weinberg assumption simultaneously. For example, a student might predict that migration could prevent fixation even in a small population. A student would then run a number of simulations of such a hypothesis and compare the results to the prediction. The resulting output could then be used to construct an estimate and a confidence interval to compare the simulation results to the prediction. As a follow-up, a student can make specific modifications to the original prediction and observe the effect in the simulated population. As a third step for gaining insight into the dynamics of real populations, students can investigate the effects of changes in mortality, brood size, and sex ratio on population growth and stability as well as allele and genotype frequencies.

## ○ Assessment

We assess learning outcomes from use of Fishbowl simulations in lectures, laboratory assignments, and examinations. Students are introduced to the Hardy-Weinberg theory during lecture and always participate in pencil-and-paper assignments in both lecture and lab. Students follow up on these activities with Fishbowl simulations to carry out homework assignments where questions are proposed regarding the expected results of specific violations of Hardy-Weinberg assumptions. For example, what is the observed result of selection against a particular genotype after a given number of generations, with specified initial conditions in the simulated population? In the lab, students are instructed to develop a prediction of the effect of a specific deviation of their own design from Hardy-Weinberg assumptions and must then test this prediction by producing an appropriate simulation experiment followed by a statistical test of the observed results. From lecture to lab, students are able to develop a stronger understanding of evolutionary forces at the single-gene level by first taking an introductory pencil-and-paper approach and then proposing and testing hypotheses using simulation-based experiments. The degree to which students can develop and analyze their own experimental simulations gives the instructor an indication of the degree of comprehension the students have acquired.

## ○ Conclusions

We find that the PopGen Fishbowl simulation provides students an easily accessible environment in which to experiment and analyze the effects of violations of Hardy-Weinberg assumptions in a simulated fish population. The model is freely available and does not require installation on individual computers. Students can devise their own experiments with replication and statistically analyze the results either individually or as group exercises. Using an agent-based simulation gives a more biologically realistic result than deterministic models and also allows students to efficiently examine in detail the dynamics of population genetics principles that were previously introduced in the lecture portion of the course.

## ○ Availability

The model is available at <http://faculty.etsu.edu/jonestc/Virtualecology.htm> and can be run in any browser with the Java 1.4.1 (or higher) plug-in.

## ○ Acknowledgments

Thanks to Uri Wilensky and all who created NetLogo for this environment in which a code-phobic professor can work. Thanks to Byron Van Nest for assisting in classroom implementation of the model. Thanks also to Andrea Edge, Nathan Weber, Sam Boyd, and Patrick Hawkins for classroom evaluation of the model. Finally, we thank William Leonard and three anonymous reviewers for their insightful comments on the manuscript.

## References

- Astad, D.N. (2007). *Populus: simulations of population biology*. [Online.] Available at <http://www.cbs.umn.edu/populus/>.
- Bishop, B.A. & Anderson, C.W. (1990). Student conceptions of natural selection and its role in evolution. *Journal of Research in Science Teaching*, 27, 415–427.
- Ferrari, M. & Chi, M.T.H. (1998). The nature of naïve explanations of natural selection. *International Journal of Science Education*, 20, 1231–1256.
- Latham, L.G., II & Scully, E.P. (2008). Critters! A realistic simulation for teaching evolutionary biology. *American Biology Teacher*, 70, 30–33.
- Perry, J., Meir, E., Herron, J.C., Maruca, S. & Stal, D. (2008). Evaluating two approaches to helping college students understand evolutionary tress through diagramming tasks. *CBE—Life Sciences Education*, 7, 193–201.
- Soderberg, P. & Price, F. (2003). An examination of problem-based teaching and learning in population genetics and evolution using EVOLVE, a computer simulation. *International Journal of Science Education*, 25, 35–55.
- Wilensky, U. (1999). *NetLogo*. [Online.] Evanston, IL: Center for Connected Learning and Computer-Based Modeling, Northwestern University. Available at <http://ccl.northwestern.edu/netlogo/>.

THOMAS C. JONES and THOMAS F. LAUGHLIN are faculty members in the Department of Biological Sciences, Box 70703, College of Arts and Sciences, East Tennessee State University, Johnson City, TN 37614; e-mail: [jonestc@etsu.edu](mailto:jonestc@etsu.edu) and [laughlin@etsu.edu](mailto:laughlin@etsu.edu).