

Web PopGen

(http://www.radford.edu/~rsheehy/Gen_flash/popgen/)

- Alleles are labeled A_1 and A_2 . There is no implied dominance between the alleles. The dominance relationship can be determined by genotypic fitness.
- Populations are distinguishable by color.
- When more than one population is simulated the upper graph depicts the frequency of the A_1 allele (p) and the lower graph mirrors the upper graph with the frequency of the A_2 allele (q).
- Two models of migration are allowed.
- Bottleneck effects may be simulated. Starting and ending generations determine the length of the bottleneck. The population size within the bottleneck may also be determined by the user.
- It is possible to compare the effects of different fitness values among the different populations.

© 2004 by [Bob Sheehy](#)
Radford University
Radford Virginia

Display:

The upper part of the display (light blue background) contains the input boxes for setting variables (population size, fitness, migration rates etc.). The lower portion contains two graphs. If 2 or more populations are modeled the upper graph will plot the frequency of the A_1 allele for each population and the lower graph will mirror the upper graph by plotting the frequency of the A_2 allele. If a single population is modeled the frequencies of both the A_1 and A_2 allele are plotted on the upper graph while the genotype frequencies of the Homozygote A_1A_1 , the homozygote A_2A_2 and the heterozygote A_1A_2 is plotted on the lower graph.

The simulation is started by clicking on the "Go" button and two new buttons will appear on the lower portion of the screen. The go button will be replaced by a "Reset" button which clears the plotted data and prepares for the next simulation. To the lower right is a "Continue" button which will remove the previously plotted data from the screen and continue with the simulation for the next interval of generations.

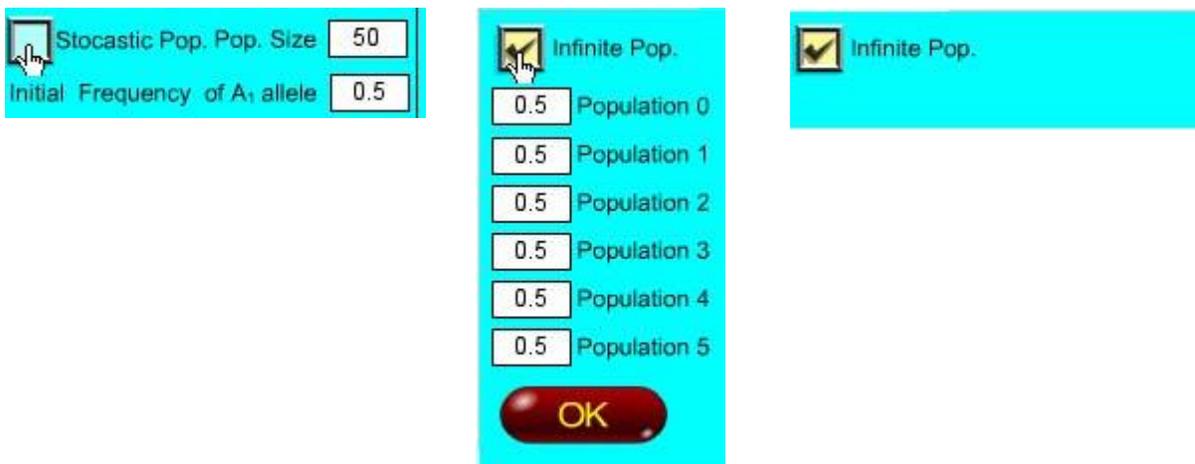
Stochastic/Infinite populations:

Simulation may be run either with a defined population size or with a theoretically infinite population. Stochastic populations are used to examine the effects of population size on genetic drift. Some other factors which affect allele frequency may become more apparent when using infinite populations.

Clicking on the square check box allows the user to toggle between stochastic and infinite populations. The number of populations being simulated will determine the number of allele frequency input boxes that are available. Below (center) is what appears for 5 populations. Each population may have a different starting allele frequency.

Population 0 is a control population. The allele frequency of Population 0 will vary with selection but migration and drift have no effect.

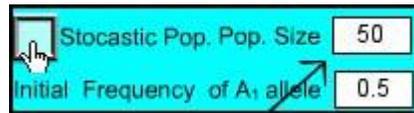
Clicking on the "OK" button will close the window causing the program to look as at right. Clicking on the check box again will return you to a simulations using stochastic populations.



Choosing infinite populations will speed the program considerably. When infinite population is chosen the initial allele frequency of each population may be set independent of the other populations. The infinite population option is unavailable if the number of different populations is greater than 5.

It is possible to toggle between Stochastic and Infinite populations during a simulation run if desired.

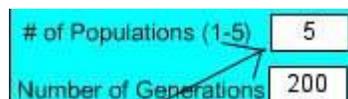
Population Size



A screenshot of a software interface showing two input fields. The top field is labeled 'Stochastic Pop. Pop. Size' and contains the value '50'. The bottom field is labeled 'Initial Frequency of A1 allele' and contains the value '0.5'. A mouse cursor is pointing at the top field, and an arrow points from the bottom field towards the right.

Population size between 5 and 250 individuals work well. The larger the population the slower the program. If one has patients and a good cup of coffee this may not be an intolerable situation to use larger populations. When running the program in stochastic mode all populations start with the same initial allele frequency. Allele frequency will drift, independently, in each population.

Number of Populations



A screenshot of a software interface showing two input fields. The top field is labeled '# of Populations (1-5)' and contains the value '5'. The bottom field is labeled 'Number of Generations' and contains the value '200'. An arrow points from the top field towards the bottom field.

1-5 populations may be simulated at a time. If 1 population is chosen then both the A_1 and A_2 allele frequencies are plotted on the upper graph and the three genotype frequencies (A_1A_1 , A_2A_2 and A_1A_2) are plotted on the lower graph. If 2 or more populations are simulated then the frequency of the A_1 allele for each population is plotted on the top graph while the frequencies of the A_2 allele for each population is plotted on the lower graph. Each population is given a different color and, in the absence of migration, each population is independent of the others.

Selection/Fitness



A screenshot of a table titled 'Fitness'. The table has three columns labeled A_1A_1 , A_1A_2 , and A_2A_2 . Each column contains the value '1'.

Fitness		
A_1A_1	A_1A_2	A_2A_2
1	1	1

Web PopGen may be used to model natural selection of alleles and genotypes. The fitness of a particular genotype may range from 0 to 1 inclusive. Fitness provides the relative survival rate of a particular genotype with the most fit being set to 1.0. For example, suppose the fitnesses for the three genotypes are $A_1A_1 = 1$, $A_1A_2 = 1$ and $A_2A_2 = 0.8$. This would indicate that fitness of the A_1 allele behaves as a dominant, That the A_1A_1 homozygote and the A_1A_2 heterozygote have an equal fitness (leave, on average, the same number of viable offspring) and that the A_2A_2 homozygote leaves, on average 80% of the number of viable offspring as either the A_1A_1 and the A_1A_2 genotypes.

The general formula for selection is:

- Average Fitness (\bar{w}):
 - $\bar{w} = p^2 * w_{11} + 2pq * w_{12} + q^2 * w_{22}$
 - where p = the frequency of the A₁ allele, q = the frequency of the A₂ allele, w₁₁ = the fitness of the A₁A₁ homozygote, w₁₂ = the relative fitness of the A₁A₂ heterozygote and w₂₂ = the relative fitness of the A₂A₂ homozygote.
- The frequencies of the various genotypes after selection:
 - Frequency of the A₁A₁ homozygote = $(p^2 * w_{11}) / \bar{w}$
 - Frequency of the A₁A₂ heterozygote = $(2pq * w_{12}) / \bar{w}$
 - Frequency of the A₂A₂ homozygote = $(q^2 * w_{22}) / \bar{w}$

Selecting Infinite population will cause a check box to appear in the Fitness box. Initially, the fitness for each genotype is the same for all populations.

<input type="checkbox"/> Stochastic Pop.	Pop. Size <input type="text" value="50"/>	# of Populations (1-5) <input type="text" value="5"/>	Fitness (Equal)		
Initial Frequency of A ₁ allele <input type="text" value="0.5"/>		Number of Generations <input type="text" value="200"/>	A ₁ A ₁	A ₁ A ₂	A ₂ A ₂
			<input type="text" value="1"/>	<input type="text" value="1"/>	<input type="text" value="1"/>

<input checked="" type="checkbox"/> Infinite Pop.		# of Populations (1-5) <input type="text" value="5"/>	Fitness (Equal)		
<input type="text" value="0.5"/> Population 0		Number of Generations <input type="text" value="200"/>	A ₁ A ₁	A ₁ A ₂	A ₂ A ₂
			<input type="text" value="1"/>	<input type="text" value="1"/>	<input type="text" value="1"/>

Initially, all populations share the same fitness values as placed in the various boxes. Clicking on the check box reveals a drop down menu which allows the user to enter different genotypic fitnesses for each of the populations.

<input checked="" type="checkbox"/> Fitness (Varied)					
	A ₁ A ₁	A ₁ A ₂	A ₂ A ₂		
	<input type="text" value="1"/>	<input type="text" value="1"/>	<input type="text" value=".9"/>	Population 0	
	<input type="text" value="1"/>	<input type="text" value=".9"/>	<input type="text" value=".9"/>	Population 1	
	<input type="text" value="1"/>	<input type="text" value="1"/>	<input type="text" value=".8"/>	Population 2	
	<input type="text" value="1"/>	<input type="text" value=".8"/>	<input type="text" value=".8"/>	Population 3	
	<input type="text" value="1"/>	<input type="text" value="1"/>	<input type="text" value=".7"/>	Population 4	
	<input type="text" value="1"/>	<input type="text" value=".7"/>	<input type="text" value=".7"/>	Population 5	
<input type="button" value="OK"/>					

After entering the desired values clicking on the "OK" button will close the dialogue box. Only the fitness values for population 0 will be visible.

<input checked="" type="checkbox"/> Infinite Pop.	# of Populations (1-5) <input type="text" value="5"/>	<input checked="" type="checkbox"/> Fitness (Varied)						
	Number of Generations <input type="text" value="200"/>	<table border="1"> <tr> <td>A_1A_1</td> <td>A_1A_2</td> <td>A_2A_2</td> </tr> <tr> <td><input type="text" value="1"/></td> <td><input type="text" value="1"/></td> <td><input type="text" value=".9"/></td> </tr> </table>	A_1A_1	A_1A_2	A_2A_2	<input type="text" value="1"/>	<input type="text" value="1"/>	<input type="text" value=".9"/>
A_1A_1	A_1A_2	A_2A_2						
<input type="text" value="1"/>	<input type="text" value="1"/>	<input type="text" value=".9"/>						

When the "GO" button is clicked, starting the simulation, a key to the various populations will appear to the right of the lower graph.

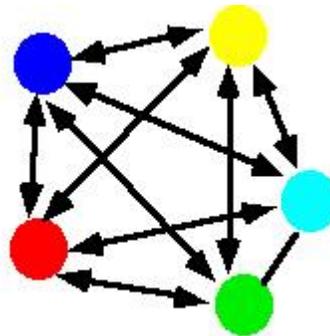
	Fitness			Population
	A_1A_1	A_1A_2	A_2A_2	
1	1	1	0.9	0 ———
1	1	0.9	0.9	1 ———
1	1	1	0.8	2 ———
1	1	0.8	0.8	3 ———
1	1	1	0.7	4 ———
1	1	0.7	0.7	5 ———

Clicking on the fitness check box will return to all populations sharing the same fitness values.

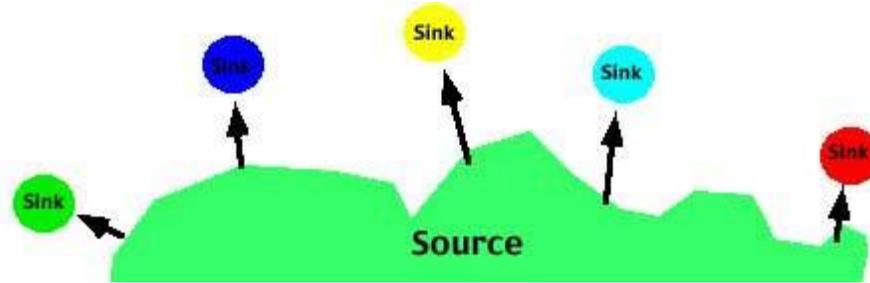
Migration

Currently two models of migration are possible.

- The Island model in which migration occurs at random among a group of small populations.



- The Source-sink or Continent-island model in which there is effectively one-way movement from a large ("source") population to a smaller, isolated ("sink") population. In this case there is no migration among the sub populations. The allele frequency(s) of the Source population may be different than the allele frequency of the sink populations.



The migration option can be chosen by clicking on the **Migration** check box. After the first click you will see the input data box asking for the rate of migration. This is a number between 0 and 1 inclusive (0 = no migration, 1 = every individual is a migrant). Clicking on the **Migration** check box again brings up the Source/Sink model where the frequency of the A₁ allele in the source population may be determined. Frequency of the A₁ allele in the source population may vary between 0 and 1 inclusive. Migration may be turned on and off during a simulation by clicking on the Migration check box while the simulation is running. This will give you an idea of the effect migration has on allele frequency among populations.

The general formula for migration is:

- Island model:
 - Average frequency of the A₁ allele (\bar{p}) is calculated across all populations.
 - $\bar{p} = (p_1 + p_2 + p_3 + \dots + p_n) / n$
 - $p_n = p_n * (1 - m) + m * p_s$
 - where p_n is the frequency of the A₁ allele in population n, m is the frequency of migrants to the population and 1-m is the frequency of individuals which do not migrate.
- Source/Sink model
 - $p'_n = p_n * (1 - m) + m * p_s$
 - where p_n is the frequency of the A₁ allele in population n, p'_n is the frequency of the A₁ allele in population n in the next generation, m is the frequency of migrants to the population, 1-m is the frequency of individuals which do not migrate and p_s is the frequency of the A₁ allele in the source population.

Mutation

Mutation Rate	
$A_1 \Rightarrow A_2$	$A_1 \Leftarrow A_2$
.0003	0

Mutation allows you to convert a proportion of one allele type to another allele type at a given frequency of mutation per allele per generation. Both forward mutations ($A_1 \rightarrow A_2$) and reverse mutations ($A_1 \leftarrow A_2$) are allowed. The frequency of forward and reverse mutations may vary.

The general formula for mutation is:

- $p' = p(1-\mu) + (1-p) * \nu$
 - where p is the frequency of the A_1 allele in population, p' is the frequency of the A_1 allele in the next generation, μ is the frequency of the forward mutation ($A_1 \rightarrow A_2$) and ν is the frequency of the reverse ($A_1 \leftarrow A_2$) mutation.

Bottle Neck

Bottle Neck?		
Start	End	BN Pop.

Bottle Neck!		
Start	End	BN Pop.
10	25	50

Bottlenecks can be simulated for stochastic populations. When the bottle neck check box is clicked three input boxes become available. Start represents the starting generation (the generation where the population crash occurs). End is the last generation of the bottleneck where the population **instantly** returns to the initial population size. The length of the bottleneck is determined by values entered into the End and Start boxes. The size of the population will remain constant during the bottleneck and is determined by the value placed in the BN Pop. box.
