EVOLVE v4.0A HELP

INTRODUCTION

EVOLVE is a computer program that allows you to experiment with evolution and to get quick results in a fashion that is impossible to do in any other way. You may control the starting population, overall population size, natural selection, pattern of inheritance and migration in a hypothetical population. By experimenting with EVOLVE you will develop:

- a better understanding of evolutionary processes and their interactions,
- a firmer grasp of some important concepts of Mendelian genetics,
- some understanding of experimental design, and
- of the use of models, plus
- an appreciation for one of the many uses of computers in biology.

BASIC EVOLUTIONARY CONCEPTS AS USED IN EVOLVE

- A **Species** can be conceived of as a metapopulation, a set of more–or–less isolated subpopulations that are capable of reproducing within & amongst themselves and which cannot reproduce with other species. EVOLVE simulates a metapopulation with one or more subpopulations.
  - Each population in a metapopulation has 1 **Gene** that affects 1 or more traits (e.g., pigments, proteins, etc.)
    - Genes may have 2 or 3 **Alleles** (usu. more in nature) symbolized by A, B, C. Alleles have
      + **Mutation Rates** (A–>B, B–>A, A–>C, C–>A, B–>C, C–>B)
      + **Patterns of inheritance**, determined by interactions between alleles and include
        – **Dominant–Recessive**: heterozygote has same phenotype as dominant homozygote, recessive phenotype is homozygous recessive. E.g., if AA = 5, AB = 5, BB = 10, A is dominant, B is recessive
        – **Incomplete Dominance**: heterozygote is in between phenotypes of homozygotes. E.g., If AA = 5, AB = 8, BB = 10, alleles are incompletely dominant.
        – **Codominance**: heterozygote displays phenotypes of both homozygotes. E.g., AA = Protein1, BB = Protein 2, AB = Proteins 1 & 2.
        – **Overdominance / Heterosis / Heterozygote superiority**: heterozygote phenotype is more extreme than either homozygote. E.g., AA = 5, AB = 13, BB =8
        – **Underdominance / Heterozygote inferiority:** heterozygote phenotype is less extreme than either homozygote. E.g., AA = 5, AB = 3, BB = 10
  - A Population consists of **Organisms.**
    - Ea. Organism has 2 copies of each gene
      + Thus, organisms may have one of 3 (or 6) **Genotypes** symbolized AA, AB & BB if there are 2 alleles; if there are 3 alleles, add AC, BC & CC.
      + Genotypes may be **Homozygous** (2 copies of same allele, i.e., AA, BB, CC) or **heterozygous** (2 different alleles, i.e., AB, AC, BC)
    - Organisms have **Phenotypes**: observable traits (e.g., pigments, proteins, structures, behaviors, etc.) produced by their genotype. Phenotypes of interest in EVOLVE include:
      + **Survival rate** (genotype's avg. proportion of young surviving to adulthood)
      + **Reproductive rate** (genotype's avg. number of young / adult)
      + **Absolute Fitness** (genotype's Survival Rate * its Reproductive rate) Values > 1 result in increasing population size, values < 1 result in population decline; they may go to extinction.
      + **Relative Fitness** (genotype's Absolute Fitness / the maximum absolute fitness) Values 0–1
      + **Migration rates** (proportion of individuals with the genotype moving to another population)
      + **Mating preferences** (proportional bias in mating of one genotype with another) Values 0–1. E.g., with *outcrossing*, individuals of a genotype prefer to mate with individuals unlike themselves. In a 2–allele population, AA individuals might proportionally mate 50% with BB, 40% with AB and 10% with AA.
- **Evolution** is defined as a change in the genetic makeup of a population, i.e., a change in allele or genotype frequencies.

HARDY–WEINBERG EQUILIBRIUM CONCEPT

- A population can be modeled as a **Gene Pool** made up of all of the alleles in a population. Gene Pools have
  - **Allele Frequencies**, the proportion of all alleles in a population that are a specific allele; e.g., the frequency of A = the number of A alleles / the total number of alleles. Traditionally, p = freq. A, q = freq. B, r = freq. C), by definition, p + q = 1 in a 2–allele population or p + q + r = 1 in a 3–allele population.
  - **Genotype frequencies,** the proportion of all genotypes that are a specific genotype; e.g., freq. of AA = number of AA / Total population size.
  - **Mating** consists of randomly sampling pairs of alleles without replacement from the gene pool to produce the young of the next generation.
- All things being equal, allele frequencies will not change and what is sometimes called the *Hardy–Weinberg Law* states that genotype frequencies will be
  $p^2 + 2pq + q^2$ in a 2–allele population or
  $p^2 + 2pq + 2pr + 2qr + q^2 + r^2$ in a 3–allele population
  This is the null case of no evolution; allele and genotype frequencies won't change and the population is in *Hardy–Weinberg equilibrium* (H–W).
- However all things are not equal. Real–word populations are subject to 5 evolutionary forces:
  - **Mutation**: H–W assumes alleles don't change, or that mutation rates from one allele to another are equal. Mutations do occur in the real world and generate the genetic variation essential for evolution to occur.
  - **Natural selection:** H–W assumes there are no differences in allele/genotype survival or reproduction. This of course is often violated and allows populations to adapt to their environments.
  - **Gene Flow:** H–W assumes the population is closed, i.e., there is no differential emigration/immigration of alleles/genotypes between populations.
  - **Genetic Drift:** H–W assumes the population is infinitely large. Any real population is finite and subsequent generations are a more–or–less accurate samples of the alleles present in the previous generation's gene pool.
  - **Sexual Selection / Non–Random mating:** H–W assumes there are no differences in probability of genotypes mating w/ ea. other.
- Violation of any of the 5 assumptions leads to changes in allele/genotype frequencies = evolution.

MODELED ORGANISMS

The organisms modeled by EVOLVE are assumed to
- Be **Diploid**: have 2 copies of each gene.
- Have **non–overlapping generations**: e.g., animals where adults mate, lay eggs or give birth and die before the young become adults, or annual plants that cross–pollinate, set seeds and die before the seeds germinate.
- **Hermaphroditic**: having male and female organs that normally mate/pollinate with another individual, but are capable of self–fertilizing if they cannot find a mate or are not pollinated.
- **Life cycle:** adults mate, produce zygotes (eggs/seed/young), die. A % of young may emigrate; migration is random between populations. Immigrants are added to resident young. Young undergo selection, those that survive are adults of next generation.

EVOLVE AS AN EVOLUTIONARY LABORATORY

Perhaps it is belaboring the obvious to say that experiments in evolutionary biology are difficult -- you can't evolve something in a semester! Nevertheless, this is an important point because it has made the approach of evolutionary biologists somewhat different from that of many other biologists. And learning about evolution is difficult because students can't "get their hands dirty" by doing experiments like those in, for example, physiology.

A common, naive view of science is that experiments are required to test hypotheses. In most scientific disciplines we note some aspect of the "real world," formulate hypotheses about major factors involved in that phenomenon and test those hypotheses with experiments. Experiments essentially are simple models of

the real world that hold most factors constant, vary one or a few factors, and observe the results. In many areas of biology, experimental design has become a sophisticated and elaborate affair of choosing organisms, equipment, statistical methods, etc.

Evolutionary biologists can apply that approach only with difficulty. We can do classical types of experiments using small organisms with short life cycles. Occasionally we can find a situation in nature that approaches a true experiment, but it is hard to coax Ma Nature into providing us with good experimental models. However, observational tests of hypotheses are quite common in evolutionary biology, as in other historical sciences like geology and astronomy.

Despite (or perhaps because of) difficulties with experimentation, biologists continue to develop models of evolutionary processes, many mathematical rather than experimental or observational. In essence, we simulate some aspect of the real world in mathematical, abstract form, and then manipulate the simulation to investigate its consequences. If the model is a good one, the consequences will clarify the real world. The Hardy-Weinberg formula and the mathematical population genetics that evolved from it are excellent examples of such models. Many of these models can be programmed into computers, which brings us to EVOLVE.

EVOLVE allows you to setup experiments that can illustrate the effects of all of the evolutionary forces on allele and genotype frequencies — Genetic Drift (population size), Natural Selection (differences in reproductive success), Mutation (change of alleles), Migration (migration of alleles between subpopulations), and Sexual Selection (non–random or preferential mating). These may be studied individually or in combination. In addition, you may allow the population size to vary within specific ranges, which allows you to see how evolution affects a key ecological variable, the intrinsic rate of population increase.

USING EVOLVE

INTRODUCTION

When you launch EVOLVE, you will see an Input Screen where you configure your experiments by selecting options and entering parameter values. When finished, click on the Submit {or Run} button. EVOLVE performs the experiment and displays a graph to visualize different aspects of the results. You may modify the data that are graphed and export graphs to PDF image files.Input parameter values and numeric results are also exported to CSV files where they may be analyzed with spreadsheet or statistical software.

INPUT SCREEN

LAB REPORT FIELDS

At the top of the input are 5 text fields that can identify output files and help you organize your thoughts about the problem at hand, the results of the experiment, and notes and thoughts. By default, it shows 5 text fields. **Range:** There is no limit on the amount of text; the fields will expand and show scroll bars.

Title Field
This is the only required field, for it becomes the labels the data and graph output files. files. A *brief* descriptive title can help you keep result files organized. E.g., "*Drift, pop 10–100*" would identify the subject and the important experimental parameter. A subsequent experiment might be titled "*Drift, pop 1000–2000"*. The other fields can be hidden or revealed by unchecking or checking the *Show lab report fields* checkbox.

Question Field
What question or problem is being studied. As you gain experience, you may want to modify this to make it more specific.

Experimental Design Field
A place for notes on how you intend to set up the experiment. Again, as you repeat more experiments you may want to modify this.

Results Field
A place to describe the data from the current experiment and list further questions.

Notes Field
A place for miscellaneous thoughts, comments, questions for additional runs, etc.

Show Lab Report Checkbox
Uncheck to hide all except Title Field, check to unhide them.

Info Button
Click to reveal a PDF file with background information and help using EVOLVE.

EXPERIMENT CONFIGURATION CONTROLS

These set up the overall configuration of the experiment, the number of alleles, duration, number of replications, and the starting population.

Number of Alleles

Select either the *2* or *3* radio buttons. Two alleles is the default. EVOLVE allows for a slightly more realistic 3 alleles, but at the cost of longer run times and the need far more inputs — 2 alleles produce 3 genotypes and 9 mating combinations, 3 produce 6 genotypes and 36 mating combinations. Most students will want to stay with 2. However, in the real world most genes have many more than 3 alleles. Students who wish to explore more complex evolutionary scenarios can try 3 alleles, e.g., to model evolution of the ABO blood groups.

Random Number seed

A number will be automatically generated in this field. You may leave it alone or change it to another number of your choice. The role of the 'seed' is a bit confusing. EVOLVE takes the "seed" and performs a series of mathematical operations to produce a series of "random" numbers that are used to, for example, randomly pick pairs of parents from the adult population, and let the survival rate vary slightly from the average values you entered in the Natural Selection panel. If you input 0.6 (60%) as the survival rate of a genotype, the rates used in 5 consecutive generations might be 61.8%, 66.3%, 53.9%, 57.5%, and 60.3%, for an average of 59.96%. If you used another seed, the 5 survival rates would be different, but would still average out to about 60%. This simulates processes such as variation in weather or in number of predators.

In summary, if you use the same seed in different experiments you can be sure that any variations in output are caused by changes in other variables. If you use different seeds with no change in other parameters, you can assess the influence of chance. In essence, this amounts to running the same experiment again. Moreover, if you want to ask your instructor about a particular run, using the same seed and inputs will produce the same exact results.

Active Evolutionary Forces

These checkboxes determine which of the evolutionary forces are active. Unchecking boxes turns off those evolutionary forces, gray out the panels' controls and prevents you from entering values, reducing the number of inputs you need to enter. Details on each force are discussed in separate sections below.

Genetic Drift
You cannot uncheck this box because every population is of finite size and thus subject to at least some genetic drift. Each generation is a random or biased *sample* of the previous generation's gene pool so the smaller the population size, the greater the effect of sampling 'errors'.

Natural Selection
Checking this box will allow you to vary survival and reproduction rates of each genotype to see how differences in fitness affect evolution. Depending on the values that are entered, you may set any pattern of

inheritance to model selection for or against alleles that are dominant, recessive, overdominananant, etc. Leaving it unchecked means that all genotypes have equal fitness.

Mutation
If this box is unchecked there will be no mutation. Checking this box allows you to set mutation rates from each allele to the other(s).

Migration
If unchecked, each subpopulation is isolated from the others and they all evolve independently. Checking this box allows you to set the amount of migration between subpopulations.

Migration may be random for all genotypes, or you may set different rates by genotype, e.g., to model populations similar to dandelions where one allele produces larger seed plumes that disperse longer distances than another.

Sexual Selection
Checking this allows you to explore the effect of non–random (preferential or biased) mating. You will enter values to determine whether genotypes tend to mate more or less often with some genotypes than they would at random — which is what happens if the box is unchecked.

The Initial Population

This section sets up the starting population for all of the subpopulations in the experimental run.

Number of Generations
Determines the number of generations EVOLVE will run before writing the CSV file and displaying the graph of the results. Start with 100 generations and adjust as needed. There is no upper limit to the number of generations, but large values (above about 10,000 — do not enter a comma) can result in long run times.

Number of Subpopulations
Running with more than 1 population allows you to get a feel for the sort of variations that can occur due to random variations in the course of evolution. A small value (5–10) is a good start; you can increase it if needed. Again, large values can increase run times.

To study migration, you must run multiple subpopulations.

Population Size
This value, along with the Initial Allele Frequencies determines the makeup of the first generation. If there are multiple subpopulations they are identical in the first generation.

EVOLVE sets no limit on the value you enter, but avoid values above 10,000. Larger values can result in very slow experimental runs. If you wish to study larger samples, simplify the experimental design by choosing a small number of subpopulations, etc.

Initial Allele Frequencies
NOTE: Because some values are calculated for you in this and some of the evolutionary forces panel, you will need to click on the *Edit* button to the right of the screen, enter values, and then click on *Apply*. Calculated values will be entered for you.

Allele frequencies will be values between 0.0 and 1.0. By default the *Enter Allele Frequencies* radio button is on. Because they sum to 1.0, you need enter only the first one (2–allele case) or two (3–alleles) values. Click on *Apply* and EVOLVE will compute the 2nd (or 3rd) frequency, as well as the number of each genotype (according to the Hardy–Weinberg equilibrium).

If you wish to start with a non–equilibrium population, click the *Enter Genotype Numbers* radio button and enter specific integers. They will need to add up to the Population Size. E.g., if you wish to see what happens to a single mutant A allele in a population of 100, enter 0 in the *AA* box 1 in *AB* and 99 in *BB*.

GENETIC DRIFT CONTROLS

Constant Population Radio Button

Genetic drift, sampling error due to small population size, is often modeled as being of a fixed–size. If you wish to do so, click the *Constant* radio button. The population size will be fixed at the *Initial Population Size*. Note, however, if the populations reproduction isn't sufficient to make up for mortality, the population can decline to extinction.

Varying Population Radio Button

However, natural populations do change in size. Natural selection in particular changes the average reproductive success of the population, what ecologists call the *intrinsic rate of population increase*.

By default EVOLVE has the *Varying* radio button selected to allow the population to increase or decrease. If the number of young produced in the population exceeds, the *Carrying Capacity* value, then the number is reduced to the smaller *Post–Crash Population Size* value. This simulates random mortality, e.g., caused by a population becoming too large for its food supply. If the populations reproduction isn't sufficient to make up for mortality, the population can decline to extinction.

NATURAL SELECTION CONTROLS

Reproduction and Survival Radio Button

Fundamentally, natural selection is the product of the survival of young to adulthood times the reproductive success of the survivors. Most evolution simulations allow students to enter *relative fitness* values that are several steps removed from the basic processes of survival and reproduction.

By default this option is checked and requires entry of two parameters for each genotype. We recommend that students start with this so they get a feel for the relationship of survival, reproduction, and fitness. Later they can change to the shortcut of the *Absolute Fitness* radio button.

Survival Rates Parameters

Survival rates may range from 0.0 for a lethal genotype to 1.0 (or 100%) for one that cannot die (obviously unrealistic).

Reproductive Rates Parameters

Reproductive rates may range from 0.0 for a sterile genotype to 10.0 for one with a high reproductive rate, an arbitrary limit.

Absolute Fitness Parameters

The *Absolute Fitness* values are grayed because they are calculated as *Survival Rate * Reproductive Rate* after you press the *Apply* button.

Values above 1.0 indicate that reproduction exceeds mortality and the genotype's offspring will more than compensate for mortality. E.g., consider 100 young of a particular genotype; if survival is 0.5, then roughly 50 will survive to adulthood. If reproduction is 2.1, then 0.5*2.1 = 1.05 and the 50 adults will produce around 105 young. If *Absolute Fitness* is less than 1.0 for all genotypes, the population will decline to extinction.

Relative Fitness Parameters

Relative fitness values are always calculated by dividing each genotype's absolute fitness by the highest absolute fitness. The relative fitness of the genotype with the highest absolute fitness will always be 1.0. Genotypes with lower values will be at a relative disadvantage and will tend to decrease in frequency. E.g., a genotype with a relative fitness of 0.75 will be only 75% as successful (will be at a 25% deficit) relative to the most successful.

Absolute Fitness Radio Button

Entering survival and reproductive rates is tedious, especially with three alleles where there are 6 genotypes for a total of 12 parameters. Consequently, more advanced students can use this option to cut their parameter entry in half by entering only *Absolute Fitness* values and clicking *Apply* button to compute the *Relative Fitnesses*.

MIGRATION CONTROLS

Migration in EVOLVE is simulated as the random movement of a proportion of the young of each genotype from one subpopulation to another. Subpopulations are assumed to be randomly distributed

geographically and movement of individuals is random, i.e., we do not consider movement of organisms along a river or chain of islands.

Genotypes Equal Radio Button

Enter one value that applies to all genotypes, e.g., 0.01 would mean 1% of each genotype moves to another population before maturity. Note that the maximum migration rate is 0.9; if 100% emigrated, there would effectively be only one population. **<QUERY— do they enter the new population as young or as adults?>**

Genotypes Vary Radio Button

Enter individual rates for each genotype. For example, rates of AA = 0.2, AB = 0.0, BB = 0.01 would simulate a dandelion–like population where the A allele is recessive and produces seeds with large plumes that blow long distances, so 20% disperse to another subpopulation; the B allele is dominant and produces seeds with small plumes and only 1% leave their original population.

SEXUAL SELECTION

Sexual selection results from the Hardy–Weinberg assumption of random mating. EVOLVE allows you to simulate this by entering proportions or probabilities of one genotype 'choosing' each of the other genotypes.

For example, assume in a species like deer or antelope that the A & B alleles are incompletely dominant and the B allele produces large antlers or horns which are preferred by females. The phenotypes would then be AA = small antlers, AB = medium, and BB = very large. EVOLVE picks females at random and the females then choose males. The mating preferences might look like the following:

	AA	AB	BB
AA Preference for:	.1	.2	.7
AB Preference for:	.1	.2	.7
BB Preference for:	.1	.2	.7

10% of the females would choose AA mates, 20% would choose AB mates and 70% would choose BBs. Note that the sum of the preferences in each row must add to 1.0 (100%). **<QUERY — Are we handling the calculation of the last genotype? I don't see an *Apply* button. Also, there should be a horizontal line between this section and the *Submit* button.>**

THE *RUN SIMULATION* BUTTON

Once all the parameters of your experimental design are entered, click the *Submit* button. EVOLVE will run your experiment and display a graph of the results. You may inspect the graph, change its axes to visualize different aspects of the results, and may export graphs and numeric data for editing, printing and further analysis (see *GRAPH CONTROLS* below).

After a run, you may change any of the input screen controls and run a similar or different experiment. The original graph window will remain and can be compared with the new graph that pops up.

GRAPH CONTROLS

In this version, EVOLVE graphs have only time (Generation number) on the horizontal axis, but you can change the vertical axis to a variety of allele and genotype frequencies (allele is the default), and numbers, e.g., Population size, number of different genotypes, etc. Click and unclick appropriate Y–Axis choices and then the *Apply* button.

If you wish to zoom in on a particular region of the graph, you can enter new coordinates in the *Y–Min*, *Y–Max*, *X–Min* and *X–Max* fields, then click the *Apply* button.

You may export graphs and numeric data. Graphs export to PNG (Portable–Network–Graphics) image files; data export to CSV (Comma–Separated–Values) files that may be opened in a spreadsheet or statistical package for more detailed analysis and graphic displays.

When you have a graph you wish to save, click the *Export Graph* button. You will see a dialog box where you must enter a name for the file, along with the .PNG suffix, and choose where in your file system to write the file.

To export data files, click the *Export Data* button, enter a file name and .CSV suffix, then choose where to save the file.

We suggest that file names be brief, but include abbreviated critical parameters and what is graphed. E.g., suppose your experiment title is "*Strong Selection for Dominant Allele*", the absolute fitnesses are 8.0, 8.0 and 1.0, and you've graphed the allele frequencies, you might use the following file name: "*Sel 4 Dom 8|1 fr A&B.png*".

A subsequent experiment with weaker selection might be exported as "*Sel 4 Dom 5|1 frA&B*".

Depending on your platform you may or may not be able to use periods, slashes, colons or other characters within file names. Use ingenuity to come up with suitable naming conventions.

EXPORTED CVS FILES

\<QUERY — need to get a look at them.\>