Evolve
Version 2.04

User's Manual

Frank Price   Hamilton College
Virginia G. Vaughan   BioQUEST Curriculum Consortium

A BioQUEST Library VII Online module published by the BioQUEST Curriculum Consortium

The BioQUEST Curriculum Consortium (1986) actively supports educators interested in the reform of undergraduate biology and engages in the collaborative development of curricula. We encourage the use of simulations, databases, and tools to construct learning environments where students are able to engage in activities like those of practicing scientists.

Email: bioquest@beloit.edu   Website: http://bioquest.org

Editorial Staff

Editor: John R. Jungck   Beloit College
Managing Editor: Ethel D. Stanley   Beloit College, BioQUEST Curriculum Consortium
Associate Editors: Sam Donovan   University of Pittsburgh
                      Stephen Everse   University of Vermont
                      Marion Fass   Beloit College
                      Margaret Waterman   Southeast Missouri State University
                      Ethel D. Stanley   Beloit College, BioQUEST Curriculum Consortium
Online Editor: Amanda Everse   Beloit College, BioQUEST Curriculum Consortium
Editorial Assistant: Sue Risseeuw   Beloit College, BioQUEST Curriculum Consortium

Editorial Board

TABLE OF CONTENTS

List of Figures

Part 1. Learning To Use Evolve

This manual has three major parts. Part 1 teaches you to use EVOLVE, Part 2 teaches you something about asking evolutionary questions through suggested exercises with EVOLVE, and Part 3 contains reference material on EVOLVE and how it relates to evolutionary biology.

An additional manual, the *Getting Started Manual*, is available. The *Getting Started Manual* contains only Part 1: Learning to Use Evolve and may be all that most students will need to begin using EVOLVE. The full *User's Manual* should, however, be available in lab or computer facility.

Preface. What You Need To Know

This manual assumes that you are familiar with operation of the Macintosh, including pointing, clicking, dragging, double-clicking, editing text, opening applications, and saving and opening documents. If not, you should learn these basic Macintosh operations before continuing with this tutorial.

In **Part 1: Learning To Use Evolve**, you will do the sample exercise in Chapter 2, "Getting Started," to get a feel for how to use EVOLVE. Chapter 3, "More Advanced Features of EVOLVE," contains three more sample exercises to give you experience with all of EVOLVE's features. You should do at least the first two exercises; the last may be omitted unless you will need to model a changing environment. You could then skim through Chapter 4, "Background," to get some perspective on population genetics and on the conceptual design of EVOLVE. Later, when you are more familiar with both EVOLVE and evolution, you may find it worthwhile to reread Chapter 4 with more care.

Part 2, Experimenting with Evolution, contains a number of sample exercises which both illustrate the capabilities of EVOLVE and demonstrate many features of evolutionary processes. Your instructor may assign some of the exercises in Part 2, Chapters 5-7, or may have you do others of his or her design.

Chapter 5, "Elementary Exercises," examines Hardy-Weinberg equilibrium and four evolutionary forces (selection, mutation, drift, and gene flow) singly, then in combination. (If you don't already know some of these technical words, there is a Glossary at the end of this manual.) The initial exercises are spelled out in detail, and subsequent exercises leave more and more to be filled in by students. The intent of this chapter is to give you a rather qualitative exposure to evolution. Although you will be looking at numerical measures of allele and genotype frequencies, we don't expect in-depth comparisons of EVOLVE's output with theoretical predictions. This chapter will be the meat of EVOLVE for the majority of students up through college undergraduates.

Chapters 6 and 7, "Intermediate Exercises" and "Advanced Exercises," are rather brief, for they are intended to point the way to additional work for advanced undergraduates and graduate students. Here the intent is to illustrate how to use EVOLVE as a microcosm to provide experimental data that may be used to test quantitatively predictions generated by equations. Although EVOLVE is a rather simplistic model, it can rapidly generate data which can be compared with theoretical predictions. Again, the exercises are of gradually increasing difficulty and assume increasingly mathematical background. The last exercise in Chapter 7 can be a sobering experience, for it brings home the enormous difficulty of "proving" what is happening in a given evolutionary situation.

Part 3, Further Considerations, contains reference information on EVOLVE's menus and screen displays, setting up experiments, and some more advanced topics. The two chapters, "Setting Up Evolutionary Experiments" and "Theoretical Notes," should be used as references when you have questions about using the program. Beginning

students may wish to read this material, but may find some of it heavy going. More advanced students will find it useful even if it is not assigned.

Chapter 1.   Introduction

EVOLVE is a computer program that allows you to experiment with evolution and to get quick results that are 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,

   • a greater understanding of experimental design,

   • a greater understanding of the use of models, and

   • an appreciation for one of the many uses of computers in biology.

EVOLVE provides abundant opportunities to practice posing questions about evolution and to try various strategies to answer those questions. It also provides data and graphs that help answer the questions as well as help persuade others of the value of those answers.

Real experiments in evolutionary biology are difficult — you just cannot evolve something in a semester or even a lifetime! This point deserves emphasis, because it requires the approach of evolutionary biologists to be somewhat different from that of many other biologists. Even learning about evolution is difficult because students cannot "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. In essence, experiments are simple models we construct of the real world that hold most factors constant. We then 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 such things as organisms, equipment, and statistical methods.

Evolutionary biologists can apply that approach only with difficulty. We can test some hypotheses 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. It is especially difficult to test such hypotheses as, "birds evolved from dinosaurs."

An alternative approach to experimental hypothesis testing is observational testing. If we hypothesize that birds evolved from dinosaurs, then we might predict the existence of fossils that show a mixture of bird-like and dinosaur-like characteristics. Such

observational tests of hypotheses are quite common in evolutionary biology and other historical sciences such as geology and astronomy. However, some aspects of evolutionary biology cannot be studied by observation or by experiment.

Despite (or perhaps because of) such difficulties, biologists continue to develop models of evolutionary processes, but many of their models are conceptual, often 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 clarify the real world and even suggest observational or experimental tests. The Hardy-Weinberg formula and the mathematical population genetics that evolved from it are excellent examples of such models. (See Chapter 8 for a detailed discussion of the Hardy-Weinberg concept.) Many of these models can be programmed into computers, which brings us to EVOLVE.

Chapter 2. Getting Started

Introduction

Using EVOLVE is easy and will rapidly become second nature, In this chapter we will take you through a simple experiment to give you a feel for the way the program works. In this manual, special keys on the Apple keyboard are shown by words or symbols enclosed in square brackets, [and]. For example, [tab] indicates the tab key, [return] denotes the key labeled as such, Directions are marked with a ☛.

Three parts of this manual are more important than their page numbers suggest — you should look them over soon and consult them when you have questions: Chapter 8 (Program Notes and Setting up Evolutionary Experiments) will be useful in understanding EVOLVE itself and when you need to design experiments. Chapter 9 (Theoretical Notes) contains background material on Hardy-Weinberg and population genetics, conceptual models, and the concepts of fitness and selection. Also, there is a Glossary of terms used in this manual — use it.

☛ **Start EVOLVE:** If the program is not already running, insert the EVOLVE disk, and turn on the computer. Open the disk icon, if it is closed, and double-click on the EVOLVE icon. **While the program is starting, continue reading.**

The Help System

If you have questions or problems about using EVOLVE, help is available from within the program. There are three ways to get help.

1. Information about inactive (gray) menu items: If you do not understand why a particular menu item is not available, you can click on an inactive item. A small window will appear giving information on the item and indicating why it cannot be chosen at this time. For example, choosing the Cut item in the **Edit** menu displays the following message: *"This option is only active when some text is selected."*

2. General information: For help with a particular menu item or about a window on the screen, you can enter Help mode by holding down the Command (or) key while typing a question mark. (If you have a [help] key on your keyboard this will have the same effect.) The cursor will change to a question mark. Clicking on any item in a menu

(all items are enabled in this mode) gives general information about that item, including what would happen if that item were chosen.

Similarly, clicking anywhere in the front-most window while in Help mode will display some general information about that window. This information includes a description of the contents of the window and explains the function of any buttons or other controls on the window.

Help mode will automatically turn off once you have clicked on an item.

3. To browse through all of the information available about EVOLVE: Choose Help with Evolve under the menu on the far left. This allows you to choose from a list of topics and display information on the topic that interests you. Buttons at the bottom of the window let you move to additional topics in the list.

Exercise 1: A Simple Experiment with Natural Selection.

In this exercise we will do an experiment to show the fate of an advantageous recessive allele that initially is uncommon. It is often useful to envision a specific example to make an abstract exercise concrete, so think of modeling a situation where a very large flock of migrating birds was blown onto an island and colonized it.

EVOLVE models evolution of one gene having two alleles, • and ◊, and thus three genotypes, ••, •◊ and ◊◊. If individuals with the •• genotype lay an average of 8 eggs, while the other genotypes (•◊ and ◊◊) lay an average of 5 eggs, the • allele is an advantageous recessive. The fact that the •• genotype produces more offspring means that it is likely to be favored by natural selection, and thus to have an advantage over the other genotypes.

For the present, we will define *dominant* alleles as ones that produce their full phenotypic effect even in heterozygotes. *Recessive* alleles have their full effect only in homozygotes. (We will discuss these and other definitions in detail later; our purpose here is to teach you how to use EVOLVE, not genetics or evolution.) Thus, because the •◊ genotype has the same phenotype as the ◊◊ homozygote, the ◊ allele is dominant and the • allele is recessive. This is like Mendel's peas — he called the allele for yellow seeds dominant over the recessive green because heterozygotes looked just like seeds homozygous for yellow. The recessive green trait appeared only in seeds that were homozygous for the green allele.

Assume further that the birds' alleles have no effect on survival, and all genotypes have a 22% chance of surviving to produce young. To ensure that selection is the only evolutionary force operating on the population, we need a large population and no gene flow. We will start a population of 8000 organisms with a • allele frequency of 5%. (If you don't know about genetic equilibrium, don't worry; it will come later.) To make the • allele recessive and advantageous, we will give the •• homozygotes a higher reproductive rate, 8, than the other genotypes, both of which will be 5.

Setting up the Experiment

☞ **Look for a window similar to Figure 2-1 below.** If you do not see one, pull down the **File** menu and select New Problem, then click on the New button.
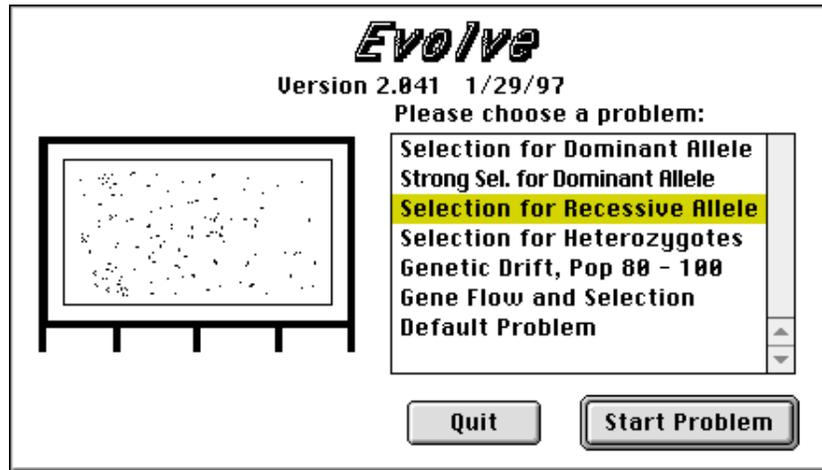
Figure 2-1. Problem Selection Window. This window appears when you start EVOLVE or when you select the New Problem command from the **File** menu.

☞ **Click on Sel. for recessive allele, and then Start Problem.** You should see the following window:
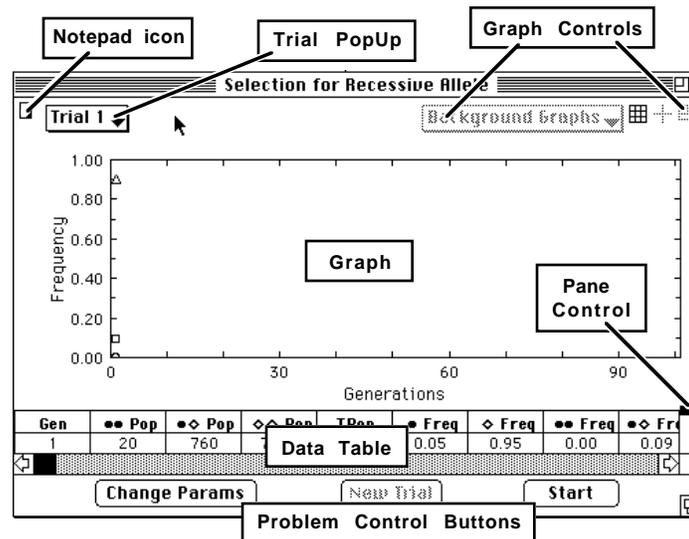
Figure 2-2. Problem Summary Window. This window allows you to run and stop experiments, and display their results.

We will refer to the labels in this picture throughout this tutorial.

☞ **Click on the Change Params button.** You will see Figure 2-3.

Figure 2-3. Parameter Input Window. This window allows you to change any of the variables (parameters) that control a trial

The Change Parameter window allows you to enter and modify the parameters, or variables, that control a trial. Parameters are contained in "edit boxes." Information in the window that is not visible or not enclosed in an edit box cannot be changed.

Initially, the window contains sample **default** values. The initial values depend on the problem selected in the Problem Selection window (Figure 2-1). You will change the values to set up your experiments. Proper choice of values for these parameters will allow you to establish a population and to determine:
- the patterns of inheritance (such as dominance-recessiveness),
- natural selection,
- gene flow, and
- population size.

These are the heart of EVOLVE's model of evolution. The other items in this window do not affect any of the evolutionary forces, but do determine how the computer manages the experiment — such things as its duration and its title.

Note that the experiment's title box ("Trial 1") is selected. The title is to remind you of what the experiment is about.

Since many of EVOLVE's parameters interact, they are edited in groups. A typical Macintosh edit box surrounds editable values.

When you first click on a box to edit its value, boxes that are not part of that group turn gray to indicate that they cannot be edited. Members of the same group remain white and may be edited normally. When you are finished with a particular group, click the Update button.

☞ **Make sure that "Trial 1" in the Title rectangle is highlighted.** If you clicked elsewhere by mistake and Trial 1 is not highlighted, use the mouse to select it before typing. Type: Trial 1- Sel For Recessive . If you make a typing error, use the [delete], [➡], and [⬅] keys and correct the error. Click on the Update button.

Note that when you update the window, the gray areas disappear. You should always be sure to enter a brief, descriptive title for your experiments. You will eventually accumulate many different experiments and their titles will help you keep them straight.

There is much to discuss about this window, but we will come back to it later. Now you should actually run the experiment.

☞ **Click on the Done button** to return to the summary window.

Doing the Experiment

☞ **Click on the Start button** in the lower right-hand corner of the window. Note that it changes to Stop.

 While the experiment is running, the window will develop into something resembling Figure 2-4.
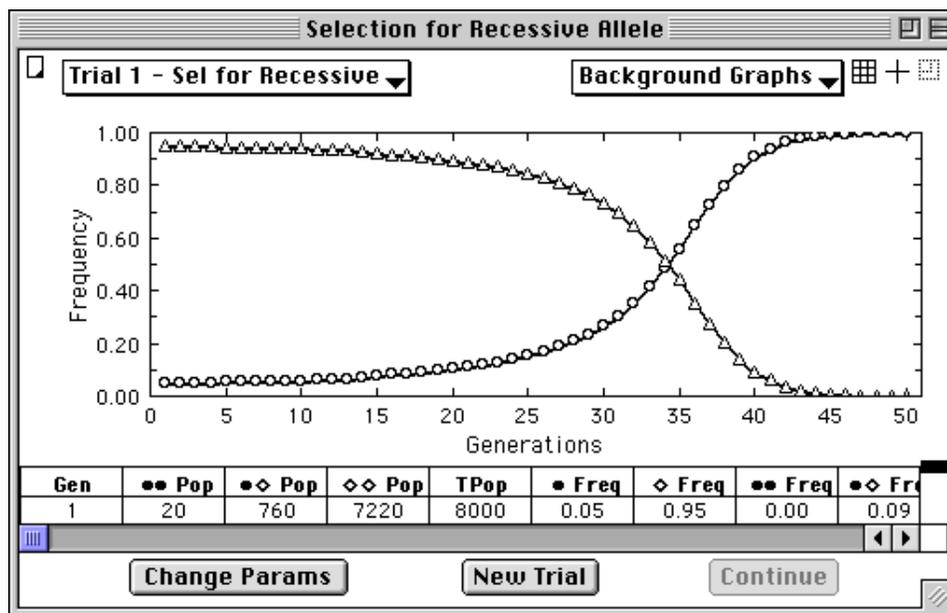
Figure 2-4. Summary Window after Trial 1. When EVOLVE
has finished the experiment for Trial 1, your summary window
should look like this.

Note that the display in Figure 2-4 shows a graph of the frequencies of each allele on the vertical axis, with time on the horizontal axis. There is also a table showing numeric summaries of each generation. As the experiment proceeds, the graph is updated,

If you wish to pause for thought or discussion, or if you realize you have made a mistake, press the Stop button. It changes to Continue, If you wish to display other graphs of current results, or start another experiment, you may do so, If you decide to continue a stopped experiment, click on the Continue button (you may not, however, exceed the maximum number of generations).

Before EVOLVE runs an experiment, you should try to predict what will happen on the screen. Predicting the results of experiments is an essential part of science and you should practice it whenever you can.

Looking at Results

Once the experiment is finished, you can examine the results in a variety of ways.

☛ **Click on the Change Params button.**

This returns to the parameters window (see Figure 2-3), which shows the parameters you entered before you ran the trial. Note that the initial population parameters cannot be changed (they do not have edit boxes around them) after the trial has been started. If you forget the experiment's initial inputs, you can go back to this window and refresh your memory. You may also change some of the values and continue the trial.

☛ **Click on the Graph menu and hold the mouse button down.** Note that there are check marks beside the frequency of each allele plotted on the graph.

☛ **Select Frequency • Allele; it is** removed from the graph. **Select it again** and it reappears.

☛ **Select Frequency of •◊ Genotype .** This adds the check mark and the line on the graph. The menu and graph should look like Figure 2-5.
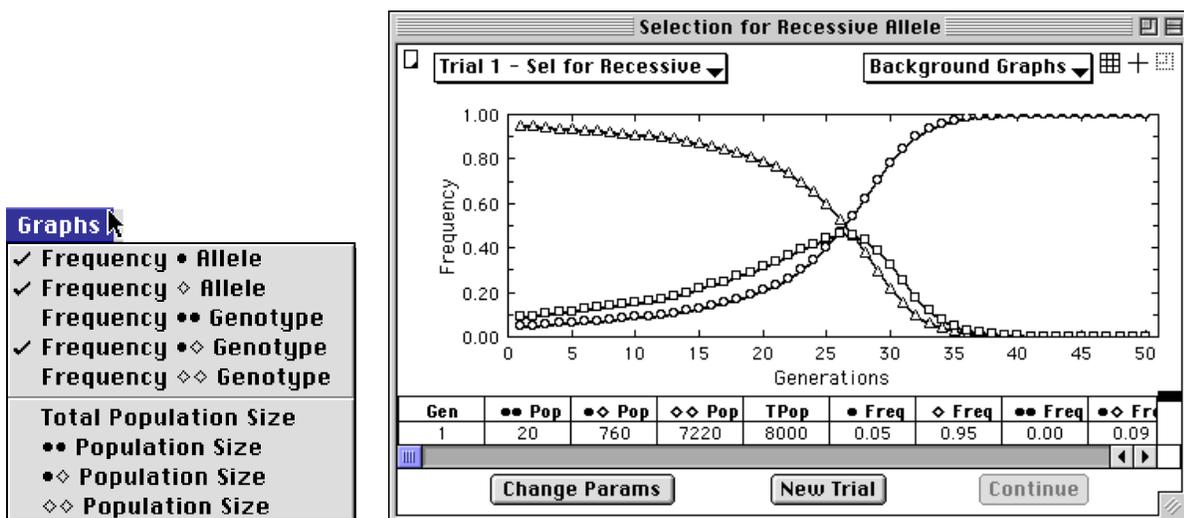
Figure 2-5. Menu and Graph of allele and heterozygote genotype frequencies vs time.

There is a considerable amount of information to be gleaned from comparison of graphs such as this and it will take you some time to become proficient in extracting all that there is to be seen. At this time we will just mention a few significant points.

☞ **Point to the •Freq heading of the data table and hold the mouse button down.** The line on the graph corresponding to the frequency of • allele remains black, and the others become gray. **Point and press on the other column headings in turn (scroll the window to the right, or enlarge the window to see hidden columns).** Each time, the appropriate line appears black and the others dim to gray. If a variable is not on the graph, it appears as long as the mouse button is down.

☞ **Click slowly three times on the gridded box, ⊞, next to the Background Graphs pull–down menu.** This toggles grid marks on the graph.

You should be able to describe what happened in graphs such as these. A full description of this graph might include the following points:
•   The frequency of the recessive • allele was initially low (about 0.05, or 5% in generation 1) and climbed relatively slowly.
•   After roughly generation 20, the • allele reached 20% allele frequency.
•   It then rose more rapidly, becoming the most common allele after about generation 30.
•   After about generation 35, its rate of increase of the • allele slowed abruptly as its frequency approached 100%.
•   The dominant, but disadvantageous, ◊ allele followed a mirror-image path and was almost extinct in generation 50.
•   The frequency of the heterozygotes peaked just below 50% when the alleles were equally common.

Note that your results will differ slightly from those shown in this manual. EVOLVE incorporates the randomness of evolution, hence each run will differ from others. The degree of difference will depend on a number of factors and will be discussed later. For this example, you should see little difference.

You can't see precise details on graphs such as this, but you can scroll back through the Results table in the Summary Window if you want to look at specific generations or frequencies. Here's how:

☞ **Click on the pane control, hold the mouse button down, and drag the black rectangle up.** The data table is dragged up revealing a scroll bar. The graph becomes shorter to accommodate the table.

You can now scroll through the data table to see exact figures for each generation.

☞ **Scroll vertically until you can see the generations after 30.**

Now you can see when the heterozygote frequency (•◊ Freq) peaked (roughly 46% around generations 35 in our trial). There are theoretical, mathematical reasons why the heterozygotes peaked, and the homozygotes crossed, at the time the allele frequencies

crossed. They are worth discussing with fellow students or your instructor, but we will not elaborate here.

☛ **Select Total Population from the Graph menu and remove Frequency ◊ Allele and Frequency •◊ Genotype.** Note that only one axis label can be displayed. **Select Frequency • Allele** to remove it. **Click on the + next to the grid toggle.** The cursor changes to a +. **Move the cursor to the 7000 line near, but not on, the vertical axis. Press the mouse button down and drag a rectangle around the interesting part of the graph.** (IMPORTANT: the rectangle must not touch the outside lines of the graph area. If it does, no line will appear on the new graph.) The resulting graph is shown in Figure 2-6:
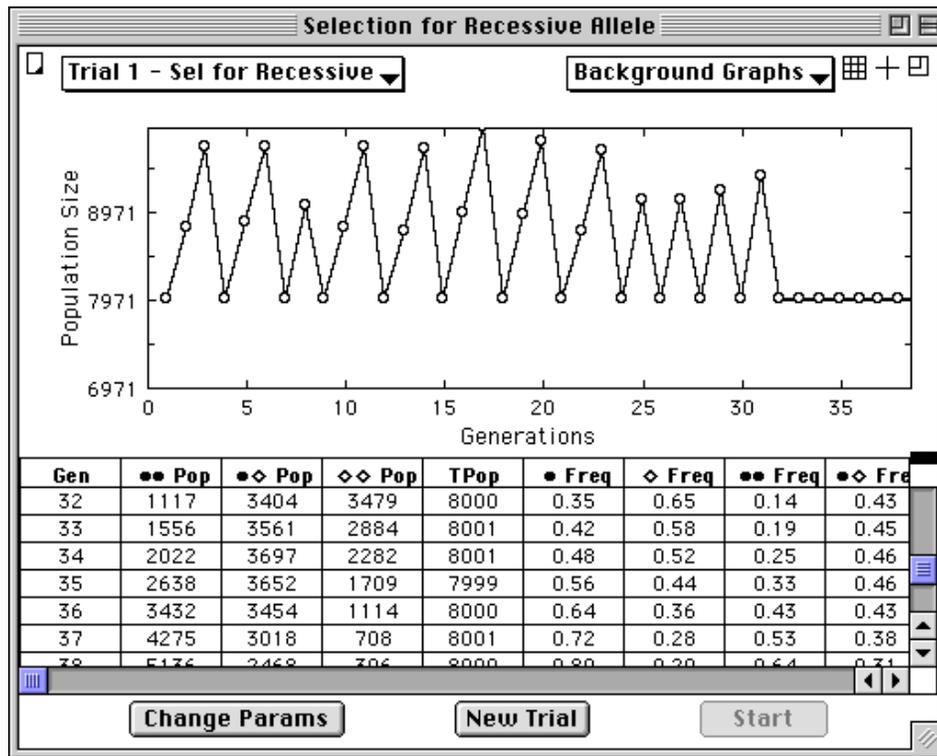
Figure 2-6. Total population size vs time.

Observe that the population size started at 8000 and climbed to over 9500 in the 3rd generation, then fell to the 8000 level in the 4th. Recall that the maximum population size in the parameter window was 9999. During the experiment, the population of adults in generation 4 was greater than 9999, so EVOLVE "killed" enough to bring the numbers down to 8000, the "post-crash" size. The population continued this saw-tooth pattern of growth and crash, with the crashes becoming more frequent so that after generation 24 it was crashing every other generation. After generation 32, the population growth rate was so great that it grew to over 9999 and crashed in every generation.

☛ **Click slowly three times on the ▢ box.** The graph zooms out to its original scale, then in, then out again.

Examine any of the various graphs as much as you wish. If you have time, and have read through Chapter 3, you may go on to the exercises in Chapter 4. If not, leave EVOLVE as follows.

Leaving EVOLVE

☞ **Select Quit from the File Menu.** EVOLVE will quit and return you to the Mac's desktop. You will be given an opportunity to save your data. **Click on the Quit button.**

This completes our first quick look at EVOLVE. If you wish to continue exploring how to use EVOLVE, turn to Chapter 3. If you wish to examine some of the background concepts that lie behind EVOLVE, jump to Chapter 4, then come back to Chapter 3.

Chapter 3. More Advanced Features of EVOLVE

Introduction: Doing Experiments with EVOLVE

In the last two chapters we introduced you to the most basic uses of EVOLVE — setting up an experiment, running it, and looking at the results — and to some background information on the conceptual design of EVOLVE's model of evolution. In this chapter we will build on those foundations, showing you some additional features of EVOLVE in the context of answering evolutionary questions by comparing a series of experiments.

Before jumping into experiments, however, it is important that you put experiments into their proper context. Experiments should not be done haphazardly; they should be done in the context of a specific question (see Figure 3-1). The question should be rather specific and you should set up a set of at least two experiments to test it. One or more experiments should be designated "controls," and used for comparisons with the other(s). You should also try to predict what the results will be in rather specific terms.
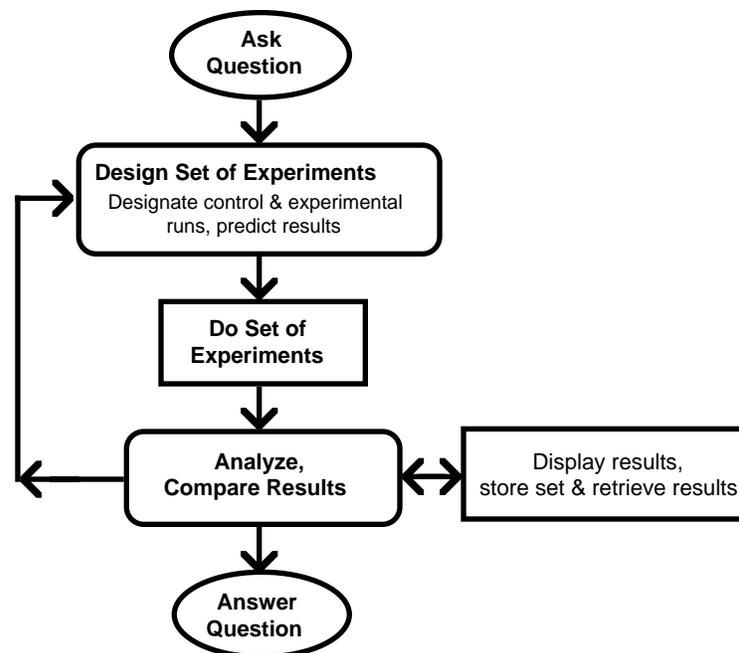
Figure 3-1. Outline of an EVOLVE session. Rectangular boxes represent actions performed using EVOLVE; rounded boxes and ovals show where user thought is involved.

After the experiments are designed, you will set up the first one and then do the experiment. Once the results are in, you may want to store them in EVOLVE's Notepad. You will then revise the evolutionary situation and do the next experiment. Comparing the results of the experiments may lead you to revise and refine your experiments or may give you enough information to answer your original question. Often, of course,

the answer to the original question will suggest additional questions and the cycle will repeat.

With this additional perspective, in Exercise 2 you will be taken through the process of answering a question about the effect of population size on evolution. In Exercise 3, you will investigate the effects of random factors on EVOLVE's results (and evolution). Finally, in Exercise 4 you will be shown how to model more complex evolutionary situations by changing the values of EVOLVE's input parameters during a run.

Exercise 2. Comparing Selection in Small and Large Populations.

An important question in evolutionary biology has been, "what are the effects of population size on the evolution of populations?" A thorough answer to this question has required decades of work by many biologists, and some aspects of the answer are still controversial. However, you can get a feel for some of the effects with a few experiments using EVOLVE.

As phrased, the question is a bit too general. Let us start by making it more specific. Since we have already done one experiment that provided data on natural selection in a population of 8000-10,000, let us answer the more limited question: "Does selection for a recessive allele proceed differently in a small population of 80-100 individuals than in a large population of 8000-10,000?" Note that this rephrasing of the question essentially completes the process of designing the experiment.

There remains only the prediction — how do you think the two experiments will compare? If you believe that selection will operate the same in the two populations and that population size has no effect, you might predict that the genotype frequency graphs of the two experiments will be identical.

We will start with the experiment from Chapter 2. You will modify it to keep the population small, run the experiment, and compare the first and second experiments. Finally, you will make a third run to observe random effects.

☞ **If you do not have EVOLVE running, start it up** just as you did in Chapter 2.

☞ **If you just restarted EVOLVE, rerun Exercise 1.** While the experiment is running, read on.

Redoing the Experiment with Variations.

To redo the first exercise with a smaller population, we will need to keep the survival and reproduction rates the same, but change the factors related to population: the initial population, maximum population, and post-crash size. We should also change the title to distinguish this from other experiments.

☛ **Click New Trial.** The graph for Trial 1 becomes gray and the trial popup changes to "Trial 2."

☛ **Click Change Parameters.** If you have reached this point directly after doing Exercise 1, the values in most of the boxes will be the same as in that experiment. If you have started EVOLVE again, then you will have to re-enter some of the values.

☛ **Change the title to a convenient title, such as "Trial 2 Rec, small."**

This reflects the fact that selection still favors the recessive allele and the population will be small (between 80 and 100).

Now we'll look more closely at EVOLVE's error checking and help system.

☛ **Click after the last zero in the 8000 of total population and press the [delete] key (or [backspace]) once to make it 800, then click on Update.**

Note that the numbers of the genotypes change to reflect the reduced population.

EVOLVE checks the values you enter and will alert you if it detects any mistakes. To see what happens. make the following "mistake."

☛ **Change the Total Population parameter to 10000, then click on Update.** You will hear a beep and a box will appear informing you that the value you entered is "out of range" and that you should check the help messages for specifics. **Click on the OK button.**

To learn the help system and find out what the possible values are for total population, do the following:

☛ **Delete one zero, click the Update button, then the Done button to return to the summary window. Select Help with EVOLVE… from the Apple (⌘) menu. Scroll the window down past the Parameters Window until you see Initial Population. Click on Initial Population, then click on the Open button.**

You will see some notes about the initial population, including the fact that the initial population cannot total more than 9000.

☛ **Click on the window's close box to return to the summary window and click on the Change Params button to return to the parameters window.**

The starting frequency of the • allele will be the same as in Exercise 1.

☞ **Now click on the Genetic Drift button to show maximum and post-crash population sizes. Change the post-crash size to 100 and the maximum population to 80.**

All of this is important because we will be comparing the results of our second run with those of the first run — the first run will be our experimental "control." If the only difference between the two experiments is size of population, then we can more easily draw valid conclusions. If there were other differences, for example, if the post-crash size was not eight-tenths of the maximum, we could not be sure that differences in results were due to differences in population size. The ratio of post-crash size to maximum population might have an effect.

☞ **Check to make sure the other parameters are the same as those in Figure 2-2 and fix any that are not. Click on the Done button, then the Start button to start the experiment.** As the lines march across the graph, try to predict what will happen next.

☞ **When the experiment is finished, compare the allele frequency graphs** from the two exercises.

Notice that while the trace of the small population (in black) is more jagged than the gray lines of the large population, it generally follows the same pattern of relatively gradual change, then rapid change. Your own graph may differ significantly from that of the 1st trial.

☞ **Compare plots of genotype frequency for both experiments.**

Again, the two graphs are quite similar in overall shape, although the small population's curves fluctuate a good deal more. We should note that there are more sophisticated ways to compare these curves quantitatively, and if you are an advanced student your instructor may have you make such comparisons. However, our purpose here is to accustom you to comparing different runs, not statistical curve fitting, so we will make only qualitative comparisons.

☞ **Plot total population size.**

Compare the two populations. You will need to zoom in or change the front-most trial in the Trial popup to see the second one.

☞ **Plot number of •• homozygotes and number of ◊◊ homozygotes**

Notice that there are periods early in the experiment when the •• genotype did not occur in the population. Not until later did the •• genotype return for keeps. Similarly, the ◊◊ homozygotes often decline to zero, then reappear before disappearing for good. Homozygotes, after all, may be generated by matings between heterozygotes, even when the rare allele is too rare for significant numbers of homozygotes to mate with each other and reproduce.

☞ **Do a 3rd trial with maximum pop = 50 and post-crash = 40.** How does this differ from the previous 2 experiments?

☞ **Click the *New Trial* button, then *Continue*. Do this 5 or 6 times.** This should give you a better idea of the effects of randomness in a small population.

Moving Results to Paper and Elsewhere

A major aspect of BioQUEST (and of science in general) involves persuading peers of our conclusions. There are several ways to move EVOLVE's results into other programs and to paper. Not only can you obtain printed results, "hardcopy" in computer jargon, but you can use EVOLVE's results as input to other programs such as spreadsheets, statistics programs, and especially word processors.

Some of you may not have a printer connected to your computer. If so, don't worry, you can do the vast majority of your work without one. If you really need a summary or copies of graphs and don't have a printer, you may save screens to disk files, take the disk to a computer that does have a printer, and print your results there.

If you do have a printer, the Notepad is more convenient. Here is the procedure for using and printing the notepad.

☞ **Select Copy Window Graph from the Edit menu.** The graph is copied to the clipboard.

☞ **Click on the small icon in the upper left-hand corner of the summary window.** This opens EVOLVE's Notepad window where you can paste the contents of the clipboard and can also type in your own notes.

☞ **Select Paste from the Edit menu.** The graph appears in the Notepad Window.

☞ **Type a few words commenting on the graph.** The text appears and can be edited in standard Macintosh fashion.

☞ **Make sure you have used the Chooser to select a printer and that the printer is turned on. Select Print Notepad from the File menu.** The contents of the notepad are printed.

While the notepad is convenient, this version of EVOLVE cannot save it when you quit the program. You can copy and paste data into word processors, spreadsheets, and statistical packages with the Copy Window Data command in the Edit menu (it cannot be pasted into the Notepad). Graphics can be copied into word processors and programs that edit graphics. You may also use the Macintosh Scrapbook to save graphics and data, then paste them into other programs.

Finally, the Save Window Data… command in the **File** menu will save the entire contents of the data table to a text file that may be opened by a spreadsheet or statistical package.

EXERCISE 3. Comparing Runs with Different Random Numbers

This exercise is designed to provide an illustration of how randomness may affect evolution. Any experiment is subject to variations, especially experiments with small populations. We will rerun Trial 2, a second time to look at some random effects.

Before we do that, we need to remove the clutter of Trial 1.

☞ **Point to the Background Graphs box and press the mouse button.** A list of trials pops up. Trial 1 appears with a check beside it. **Select "Trial 1"** and it is unchecked and its lines vanish from the graph.

☞ **Click on the New Trial button and the Change Parameters button.** The Trial number changes and the graph lines become gray. All of the parameter values except title remain the same.

☞ **Click on the Seed button.** A number appears that is the "random number seed" for this trial. The computer uses it to simulate random mating, the effects of weather, and other "random" factors. EVOLVE will pick a different seed for each trial, or you can type a seed in if you wish. 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. A final point — if you have a trial that you wish to repeat exactly (e.g., to show your instructor), copy the seed along with the other parameter values. A thorough understanding of the seed is not necessary for you to use EVOLVE, but a more thorough discussion is available in Chapter 8 of the User's Manual.

☞ **Click on the Done button, then the Start button to run the new trial.**

Were the results what you expected? This phenomenon of random fluctuations of genotype and allele frequencies in small populations is an important, controversial evolutionary force called "genetic drift." You should run additional trials to get a feel for the variability of results. Obviously, chance can have a significant influence on evolution.

EXERCISE 4. Changing the Evolutionary Situation During a Run

One of the simplifications often made in modeling evolution is to assume that the evolutionary forces are constant, that is, the environment doesn't change. Obviously, this is a gross oversimplification and EVOLVE will let you get around it by changing data values during a run. In this final exercise of our tutorial, you will see how to do this.

Suppose you wished to simulate a drastic drop in population size, such as would occur if there was a catastrophe like a flood that killed most of a population and reduced their food supply for a couple of years. This would simulate what is sometime called the "bottleneck effect." In this scenario, 30 individuals are assumed to survive a disaster from a large population having both alleles in equal abundance. The alleles are assumed to be selectively neutral.

☞ **Set up an experiment with the following data values** (note that only 39 generations are to be done) and then run the experiment:

<div align="center">

Title: Ex. 4 - The Bottleneck Effect

</div>

	• •	• ◊	◊ ◊
Post-crash size:	8000		
Run to generation:	40		
	Genotypes		
Initial population:	2000	4000	2000
Survival rates:	22	22	22
Reproductive rates:	5	5	5
Immigrat., emigrat.:	0	0	0

☞ **When the 40 generations are finished, click on the Change Parameters button and revise the parameters as follows:**

Carrying capacity: 50
Post-crash size: 30
Run to generation: 43

☞ **When EVOLVE stops in generation 43, go back and return the variables to the following:**

Carrying capacity: 9999
Post-crash size: 8000
Run to generation: 100

☞ **Compare the graphs of allele frequency and total population size.**

The allele frequencies changed relatively little during the first 40 generations when the population cycles between 8000 and 10,000; during the 3-generation crash, however, the

allele frequencies "drifted" away from 0.50. Subsequently, as the population grew back, the generation-to-generation fluctuations tapered off.

The final population was different from the original because 1) the founders of the "new" larger population were a small sample from the original large population and 2) while the population was rebuilding, drift continued to operate, albeit to a lesser degree as the population size increased.

This completes our tutorial on EVOLVE. We hope you have enjoyed learning to play the game and will use it extensively enough to get a good feel for the interaction of evolutionary forces. The next chapters list a series of exercises that will help you use EVOLVE to explore a variety of evolutionary phenomena. The last two chapters contain additional information on setting up experiments and on population genetics; don't overlook them, for they provide more guidance on evolution and on EVOLVE.

Chapter 4. Background

Now that you have some feel for the way EVOLVE works, we must take time to go over some fundamental concepts and background so you can make more effective use of it. As mentioned in the Introduction, simulations such as EVOLVE have a large mathematical component, and you should eventually work through the material in Chapters 7 and 8. However, EVOLVE can be used effectively at an introductory level without such details, for it embodies an intuitively simple, yet realistic, model of evolution. The purpose of this chapter is to give you a better intuitive feel for EVOLVE and some of the fundamental genetic and evolutionary concepts behind it.

Also, while many evolutionary situations may be simulated with EVOLVE, it has limitations, and you must understand the nature of the hypothetical population. The next sections discuss the genetic concepts and hypothetical organisms on which EVOLVE is built; that is, the assumptions inherent in the program.

The Hypothetical Organisms

EVOLVE is not based on any particular animal or plant, but on a hypothetical organism with several characteristics that make it useful for the sorts of experiments you will be doing. The organisms live in a discrete habitat such as an island, lake, or mountain top, separated from other patches of habitat, and form a single local population within which mating is random. The individuals in the population are diploid hermaphrodites and produce both eggs and sperm. Each adult normally mates with (or is pollinated by) one other adult at random and both produce offspring. If there is an odd number of adults in the population, the last individual fertilizes itself. While this characteristic of the model population may seem odd, it avoids complications of sex-ratios and mating patterns. Actually, this sort of population is fairly common among plants and even exists in a few animals.

The life cycle of the hypothetical organism is a simple one (see Figure 4-1 for an outline of the life cycle as it relates to EVOLVE's simulation). During the short breeding season, all adults mate, produce offspring, and die. All members of the next generation hatch (are born, released as seeds, or whatever) during a short time. The young then mature over a period of time during which they may die or emigrate (fly, walk, blow or be carried away). At the end of the juvenile period all surviving individuals become adults and some additional adults may immigrate from surrounding populations, or adults may emigrate, leaving the population. You may find it convenient to think of the organism as having a one-year life cycle and generation time. You may specify any or all of the survival, reproduction, emigration, and immigration parameters for each of the three genotypes.
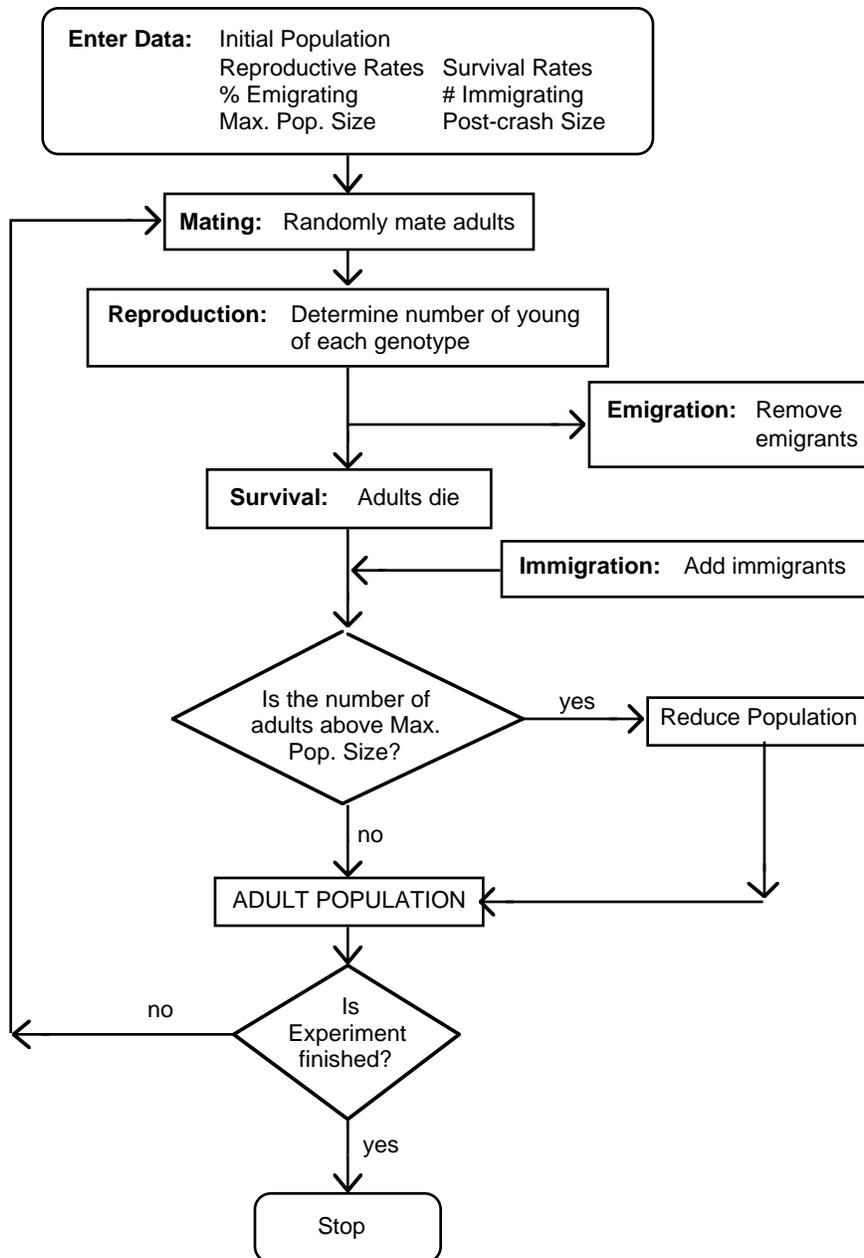
Figure 4-1. Outline of EVOLVE's simulation and the hypothetical life cycle.

Given the above life cycle, the population usually will either grow or decline depending on the values you give EVOLVE. If the product of percent survival and number of young is greater than 100%, the population will tend to increase in numbers. In order to control the size of the population, you may specify upper and lower bounds to it. If the population of adults exceeds the upper limit, the number of adults is randomly reduced to the lower level. The numbers of each of the genotypes are reduced proportionately, so a population "crash" will not directly affect genotype or allele frequencies. If survival and reproductive rates result in a negative growth rate (their product is less than 100%), then the population will decline to extinction, regardless of the population size limits. In

essence, then, the population may be viewed as living in a finite environment with a fixed amount of resources. If the population exceeds the resources, mortality is random with respect to genotype.

Note that given the above conditions, all individuals are identical with respect to sex and age, so the hypothetical population used by EVOLVE meets the assumptions of Hardy-Weinberg equilibrium. Mutation does not occur and mating is random. Because you may specify values for all of the other assumptions, you may design a population that fits any experiment you want to do on selection, genetic drift, and gene flow. For example, since the habitat is a discrete patch, you can make it closed (no gene flow) if you wish.

Genetic Concepts Underlying the Model

In genetic terms, EVOLVE deals with a single gene having two alleles, • and ◊, Thus there are three genotypes: ••, •◊ and ◊◊. The alleles may effect any or all of four characteristics: survival and emigration rates of juveniles, and reproductive and immigration rates of adults.

Rates of survival, reproduction, immigration, and emigration are not, strictly speaking, phenotypes of organisms in the same way that flower color is for a plant, or vestigial wings are for a fruit fly. Phenotype is classically defined as an observable characteristic of an organism, while the rates used by EVOLVE are statistical characteristics of *populations* of organisms. Nevertheless, because phenotypes influence the ability of their owners to survive and reproduce, we might think of the two alleles of our simulation as determining the probability of individuals surviving and reproducing. In a similar way, the color of a flower may influence how often insects pollinate it and hence the number of young produced. Among fruit flies, individuals with vestigial wings cannot fly and get trapped more often in sticky food, hence their survival rate is lower than normal flies. You will find it easier to use EVOLVE if you make up phenotypes appropriate to the question you are studying and attach them to the genotypes as we did in Chapter 2.

For the purposes of using EVOLVE, use the following definitions to determine the pattern of inheritance.

Dominant alleles produce their full phenotypic effect even in heterozygous condition, *Recessive* alleles have their full effect only when in homozygous condition. Thus, if the •◊ genotype has the same phenotype as the ◊◊ homozygote, then the ◊ allele is dominant and the • allele is recessive.

Incomplete dominance occurs when the phenotype of the heterozygote is intermediate between the phenotypes of the two homozygotes. Reproductive rates of 4, 6, and 8 for the ••, •◊, and ◊◊ genotypes would simulate incomplete dominance.

Codominance occurs when both alleles produce their effects in heterozygotes. Suppose you set survival rates to 30%, 30%, and 40% and reproductive rates to 4, 5, and 5

(offspring per adult). The • allele reduces both survival rates and reproductive rates, while the ◊ allele produces higher survival and reproduction. The • allele is dominant with respect to survival, but recessive with respect to reproduction. The end result is that both alleles produce their effects in heterozygotes when you consider survival and reproduction together. This is an example of **pleiotropy** — a single gene has multiple phenotypic effects.

Overdominant alleles produce a heterozygote which is more extreme than either homozygote. If alleles are overdominant with respect to fitness, the terms **heterosis** or **heterozygote superiority** are often used. Reproductive rates of 40%, 60%, and 30% would exemplify overdominance and heterosis.

An Example

As an example of the relationship of alleles, genotypes, patterns of inheritance, and phenotypes to survival and reproductive rates, let us consider the gene for sickle-cell anemia. The s allele causes the substitution of valine for glutamic acid at position 6 of the beta chain of hemoglobin. People with the SS genotype have hemoglobin that is entirely normal; heterozygous (Ss) individuals have hemoglobin that is a mixture of normal and abnormal; those with the ss genotype have hemoglobin that is entirely of the sickling type. In this example, the fundamental effect (phenotype, if you will) of the s allele is the production of abnormal beta chains. Thus, at the level of hemoglobin phenotypes, the S and s alleles are codominant (heterozygotes show the effects of both alleles).

Hemoglobin with s beta chains has reduced solubility under low oxygen concentrations and tends to crystallize in capillaries. In homozygous ss individuals, sharp crystals grow within red blood cells, causing them to take the distorted sickle-shape that gives the disease its name. The distorted blood cells interfere with circulation and cause a variety of unpleasant symptoms that usually result in painful death before puberty. Heterozygous individuals with the Ss genotype do not typically show any of the symptoms of the disease, although their red blood cells will show some sickling under certain conditions and they are said to have "sickle-cell trait." Thus, the disease phenotype and the s allele may be regarded as recessive to the "normal" phenotype produced by the S allele.

None of the phenotypes discussed above could be input for EVOLVE. Rather, the reduced survival of ss homozygotes would be modeled by giving ◊◊ individuals a lower survival rate than •• and •◊ individuals. Again, note that **EVOLVE incorporates the statistical effects of phenotypes on reproduction and survival, not the phenotypes themselves.**

No assumptions have been made as to the mode of inheritance of the characteristics of the three genotypes used by EVOLVE. By choosing suitable parameters you may simulate any pattern of inheritance possible at a gene with two alleles. For the sickle-cell

example, setting survival rates to 80%, 80%, and 0% for the ••, •◊, and ◊◊ genotypes respectively would simulate the recessive lethal nature of the sickle-cell allele.

It should be clear by now that the distinctions between these different modes of inheritance of phenotypes are somewhat arbitrary and not always clear-cut. In using these terms, we are dealing with patterns of inheritance of phenotypes, rather than the nature and function of genes at the molecular level. Because most (probably all) genes have multiple effects, alleles may be codominant at the molecular level, dominant with respect to one phenotype, recessive with respect to another, and heterotic for a third. EVOLVE simplifies this complexity by looking at the net effects of genes on survival and reproductive rates.

Again, the sickle-cell allele is an instructive example. As discussed above, the two alleles are codominant at the molecular level, but with respect to the disease phenotype, the s allele is recessive (the S allele is dominant). The s allele would seem to be disadvantageous, and you might expect it to decline in frequency and eventually become extinct (save for new mutations). However, the situation becomes more complicated and interesting when you consider additional information.

As you may know, the s allele occurs at a high frequency (over 20%) in some populations. Research has shown that such populations have a high incidence of malaria, and that heterozygous Ss individuals, those with sickle-cell trait, have a greater tolerance to malaria than homozygous SS individuals. Here is a summary of the genotypes and their phenotypes as discussed so far:

	Genotype		
Phenotype	SS	Ss	ss
Hemoglobin	Normal	Mixture	Sickle-cell
Blood	Normal	Sickle-cell trait	Sickle-cell anemia
Malaria	Susceptible	Resistant	Resistant
Survival rate	Low	High	Essentially zero

Note that, with respect to malarial resistance, the s allele is dominant and advantageous, but with respect to hemoglobin phenotype, it is a deleterious recessive. Overall, the s allele displays heterosis (is "overdominant") with respect to survival.

You might model such a population with the following inputs to EVOLVE:

	Genotype		
Phenotype	SS	Ss	ss
Survival	Intermediate	Highest	Essentially zero
Reproduction	High	High	Essentially zero

(You might expect reproductive rates of Ss individuals to be somewhat lower than those of SS individuals, because 25% of the children of two heterozygotes would die of sickle-cell anemia. However, it appears that such parents often have more children to make up the difference.)

In an environment without malaria, however, the following would be an appropriate table:

		Genotype	
Phenotype	SS	Ss	ss
Hemoglobin	Normal	Mixture	Sickle-cell
Blood	Normal	s-c trait	Sickle-cell anemia
Malaria		Irrelevant	
Survival rate	High	High	Essentially zero

You could model such a population with the following inputs to EVOLVE:

		Genotype	
Phenotype	SS	Ss	ss
Survival	High	High	Essentially zero
Reproduction	High	High	Essentially zero

Thus, in a malaria-infested environment the sickle-cell allele is advantageous in heterozygotes and shows heterosis with respect to survival. As you will see later, alleles showing heterosis tend to remain in a population no matter how deleterious they are in homozygous condition. In a malaria-free environment the allele is recessive and deleterious with respect to survival. In the latter environment, you would be correct in expecting it to decrease in frequency.

The Nature of Evolutionary Fitness

Evolutionary biologists' use of the term *fitness* differs significantly from common usage. It might have been better if someone had invented a long Latin word for this concept, for people would be more cautious about assuming they know its meaning. Instead, people assume that evolutionary fitness is like athletic fitness — robustness, strength, endurance. Instead, we will define evolutionary fitness as the *ability of an allele or a genotype to gain representation in the next generation because of the ability of its phenotype to survive and reproduce.* Because it is critical that you understand this definition and its implications, both to understand the behavior of EVOLVE and to understand evolution, we will spend some time discussing it.

The detailed measurement of the fitness of an allele in nature is difficult, for it requires understanding genotype frequencies and the probabilities of different matings, along with survival and reproductive rates. If you work your way through the more advanced discussion of *absolute* and *relative* fitness in Chapters 7 and 8, you will come to understand this clearly. For now, note that evolutionary fitness involves both reproduction *and* survival; you cannot look at either one alone. Thus, the evolutionary fitness of a genotype, and its evolutionary fate, may have nothing to do with whether the organism is strong, swift, or "red in tooth and claw" if that organism does not reproduce. In some situations, slight, fragile individuals, even those with seeming deficiencies, may have higher fitness than normal or more robust individuals.

In caves, for example, eyes are a decided handicap, for they are potential sites of infection and require calories and nutrients to grow and maintain. Since food availability in caves is usually limited, individuals with mutations tending to reduce eye size would have higher survival rates. This could lead to higher reproductive rates, because the physiological effort saved from growing and supporting eyes could be put into egg or sperm production. Therefore, individuals with reduced eyes would have higher fitness than those with normal eyes. In time, the population would come to consist of eyeless individuals.

This view of fitness can even result in the seeming paradox of a population becoming extinct when it contains genotypes with "high" fitness. By the above definition, a genotype has high fitness if its percentage share of succeeding generations grows larger. As a hypothetical example, suppose a population consisted of two types of individuals that did not interbreed. If one type (A) has a 20% survival rate and produced 4 young, its fitness would be higher than another (B) with 15% survival and 4 young. The first type would increase in frequency relative to the second, although its absolute numbers would decline and both would become extinct. Suppose 100 young of each type were born in one generation and we followed the population for a generation or two:

Generation	No. young of type A	No. young of type B	No. adults of type A	No. adults of type B	Frequency of type in adults A	Frequency of type in adults B
1	100	100	20	15	0.57	0.43
2	80	60	16	9	0.64	0.36
3	64	36	13	5	0.70	0.30
4	51	22	10	3	0.76	0.24
5	41	13	8	2	0.81	0.19

Of 100 type A young in generation one, 20 lived to adulthood. Each of the 20 had 4 young, for a total of 80 young A's in generation two. There were only 60 young B's in generation two. As you can see, the frequency of type A increased even though its actual numbers declined. You will see an example of this situation in the exercises. Thus, evolutionary biologists distinguish absolute fitness from relative fitness (see Chapter 7).

 Now that you have learned some of the assumptions that underlie EVOLVE, you are in a better position to work with it to design experiments that will improve your understanding of our theory of evolution within populations.

Part 2. Experimenting with Evolution

These chapters are designed to give you a sampling of exercises that will enable you to more easily realize the major goals of EVOLVE: a better understanding of evolutionary processes and of how to study them.

This selection of exercises is not complete, nor will it be appropriate for all students; no one is likely to do all of them. Rather, it provides a sample of some of the ways EVOLVE may be used. Each instructor should select some of these exercises to go with others of his or her own devising. The exercises that may be done with EVOLVE span a tremendous range of evolutionary situations. But more than that, answers can be derived using methods of varying degrees of sophistication. The grouping of exercises into three chapters should give students some feeling for the open-ended nature of evolution and of computer models.

Chapter 5 is a relatively intuitive, qualitative examination of Hardy-Weinberg equilibrium, selection, the fate of mutations, drift, and gene flow, along with combinations of drift with selection and gene flow with selection. In the initial exercises, we provide all inputs to EVOLVE and in succeeding exercises more and more of the values must be determined by the students. This gradually more difficult series of assignments is suitable for high school and freshman or sophomore college students.

Chapter 6 takes a more sophisticated, quantitative approach by having students consider absolute and relative fitness coefficients in explaining EVOLVE's results, and the effects of evolution on mean fitness of a population and its growth rate (what ecologists call "intrinsic rate of natural increase"). This chapter concludes with an exercise aimed at getting students to use the literature to find published data on sickle-cell anemia; then they try to model evolution of the sickle-cell allele. Students also can collect "data" from EVOLVE runs and then examine more abstract plots of change in allele frequency vs allele frequency, or examine allele frequency over time in a sampling of drifting populations with and without selection. Students will begin to get a better feel for the statistical nature of evolution. These exercises are appropriate for more advanced undergraduate students.

Chapter 7 outlines several statistical approaches to the study of evolution. Plots of selection coefficients over time allow students to begin to examine the relationship of fitness and selection coefficients derived from a priori survival and reproductive rates with those that can be observed from data on changes in allele and genotype frequencies over time. EVOLVE provides "field data" from which students try to infer the pattern of selection. Students can also make statistical comparisons of EVOLVE's results with those predicted from deterministic equations of the effect of selection on allele frequency over time, or of the effects of drift on mean and variance of allele frequency in a sampling of populations. These exercises would be useful for mathematically sophisticated upperclass or graduate students.

Chapter 5. Elementary Exercises

This chapter is designed to lead you through elementary experiments with each of the evolutionary forces, singly and in combinations. The first part of the chapter reexamines fundamental concepts; the second looks at Hardy-Weinberg equilibrium in the context of setting up EVOLVE experiments and reading the results; the third provides a series of exercises that guide you through experiments with single evolutionary processes; and the final section illustrates how pairs of evolutionary forces interact.

Fundamentals

Questions 1-6 below are intended to help you test your understanding of the fundamentals of EVOLVE's simulation and do not require that you make any computer runs. We strongly recommend that you do these first six questions before you attempt any of the other exercises, because for them we will assume that you know the concepts involved with these first exercises. Your instructor may suggest that you treat this as a take-home quiz after you finish Chapter 4, and have you bring in your answers for discussion or grading.

Initial Population

1. Calculate the number of each genotype in a Hardy-Weinberg equilibrium population of 2330 individuals with an ◊ allele frequency of 0.63; write the number of each genotype in the spaces below:

No. •• = _____   No. •◊ = _____   No. ◊◊ = _____

2. Consider an initial population of 448 •• individuals, 1238 •◊ individuals, and 855 •• individuals.

a. What are the frequencies of the alleles?

Frequency of • = _____   Frequency of ◊ = _____

b. Given the allele frequencies you calculated above, determine the numbers of each of the genotypes you would expect if the population were in Hardy-Weinberg equilibrium:

No. •• = _____   No. •◊ = _____   No. ◊◊ = _____

c. Is the population in Hardy-Weinberg equilibrium? _____

Survival and Reproductive Rates

For Questions 3-6, use the following data. The tables below show possible survival and reproductive rates for five "runs" of EVOLVE. Survival rates are measured in terms of percent of each genotype surviving from birth (or hatching, germinating) to reproductive age. Reproductive rates are measured as the average number of young born per individual of each genotype.

a.

Phenotype	••	•◊	◊◊
Surviv. rate	25.0%	25.0%	20.0%
Reprod. rate	3.0	3.0	3.0

b.

Phenotype	••	•◊	◊◊
Surviv. rate	50.0%	30.0%	30.0%
Reprod. rate	2.0	3.0	3.0

c.

Phenotype	••	•◊	◊◊
Surviv. rate	40.0%	40.0%	58.0%
Reprod. rate	1.0	2.0	3.0

d.

Phenotype	••	•◊	◊◊
Surviv. rate	12.0%	18.0%	24.0%
Reprod. rate	9.0	9.0	9.0

e.

Phenotype	••	•◊	◊◊
Surviv. rate	60.0%	75.0%	65.0%
Reprod. rate	2.0	1.0	1.0

f.   None of above

3. In which of the tables is • dominant for survival rate?

a.   b.   c.   d.   e.   f.

4. In which of the tables is • recessive for reproductive rates?

a.   b.   c.   d.   e.   f.

5. In which of the tables are • and ◊ heterotic for survival rate?

a.   b.   c.   d.   e.   f.

6. In which of the tables do the alleles show incomplete dominance for reproductive rates?

a.   b.   c.   d.   e.   f.

Hardy-Weinberg Equilibrium

These next two exercises will illustrate what is required of you and give you an introduction to the whole process. In particular, note the way the parameter values are set up, and how to examine the graphs of results and the types of questions asked. Also, note that for each question we have made one or more predictions about the

outcome of our experiments with EVOLVE. Subsequent exercises are more abbreviated to encourage you to develop skill in using the program and in investigating problems.

You may wish to make several runs with each set of parameters. Note that there is to be no gene flow, nor will you need to change parameters during a run.

If you are uncertain about allele and genotype frequencies and how to calculate them, read appropriate sections of your text or in Chapter 9 of this manual. Also, if you do not understand the Hardy-Weinberg Law (better referred to as "Equilibrium"), consult your text or read Chapter 9 in this manual.

Exercise 5. How Long Does It Take To Establish Hardy-Weinberg Equilibrium Starting with a Population That Is Not in Equilibrium?

Experiment and Prediction: A population in obvious disequilibrium would be one consisting entirely of heterozygotes. There are other ways to set up such an out-of-equilibrium population; can you think of several? On the basis of what you have learned in this manual, it would be reasonable to predict that the population would reach Hardy-Weinberg equilibrium in one generation. More specifically: the allele and genotype frequencies would remain stable from generation to generation, and within one generation the genotype frequencies would approximate p^2 for ••, $2pq$ for •◊, and p^2 for ◊◊.

Experiment: Make an EVOLVE run with the following data and either print a summary of each generation, or examine the appropriate graphs with care:

Title: 5. Estab. of H-W Equilibrium

Number of generations:	10		
Carrying capacity:	9999		
Post-crash pop. size:	8000		
	••	•◊	◊◊
Initial population:	0	8000	0
Survival rates:	26	26	26
Reproductive rates:	4	4	4
Number immigrating:	0	0	0
Percent emigrating:	0	0	0

Before you do this experiment, think about what you are being asked to do.

a. What is the frequency of the ◊ allele in the initial population? _____. What should the allele frequency do from one generation to the next? _____

b. What should the genotype frequencies be in a population with the above allele frequency?   ••:_____   •◊:_____◊◊: _____

c. Why are the survival and reproductive rates equal for all genotypes?

d. Why is the maximum population set at 9999 and the post-crash size to 8000?

Results: Fill in the following data table from the data table displayed by EVOLVE. These first few exercises contain tables like these to help you learn what to look for in EVOLVE's output; we will dispense with them later. If your are experienced with spreadsheets, you can also use Copy Window Data to copy the data to the Mac's clipboard. Then Paste the results into a spreadsheet and generate the tables below with that program.

	Observed frequencies				
	Genotype frequency			Allele frequency	
Generation	••	•◊	◊◊	•	◊
Generation 1:	_____	_____	_____	_____	_____
Generation 2:	_____	_____	_____	_____	_____
Generation 3:	_____	_____	_____	_____	_____
Generation 6:	_____	_____	_____	_____	_____
Generation 7:	_____	_____	_____	_____	_____
Generation 9:	_____	_____	_____	_____	_____
Generation 10:	_____	_____	_____	_____	_____

Now, subtract the appropriate value in the earlier generation from the value in later generation; record the result (with the sign) on the appropriate blank.

	Change in frequencies				
	Genotype frequency			Allele frequency	
Generation	••	•◊	◊◊	•	◊
1 -> 2	_____	_____	_____	_____	_____
2 -> 3	_____	_____	_____	_____	_____
6 -> 7	_____	_____	_____	_____	_____
9 -> 10	_____	_____	_____	_____	_____

c. Describe the changes in genotype frequencies from one generation to another (be sure to include specific data illustrating your point):

d. Describe the changes in allele frequencies from one generation to another (be sure to include specific data illustrating your point):

e. Were the genotype and allele frequencies stable from one generation to another (be sure to give data that illustrate your point)?

f. Did the genotype and allele frequencies in generations 2-10 match your predicted values?

Conclusion:

g. Explain the changes in allele and genotype frequencies in this experiment.

h. This exercise dealt with only one form of disequilibrium: the initial population consisted entirely of heterozygotes. Would populations with other types of disequilibrium reach Hardy-Weinberg equilibrium in the same way? You may want to try some other experiments of your own devising to test your understanding. Can you think of types of disequilibria that EVOLVE *cannot* model?

Exercise 6. Set up a Population in Hardy-Weinberg Equilibrium.

In the blanks below, set up another run in which the initial frequency of the ◊ allele equals 25%, the initial genotype frequencies are in Hardy-Weinberg equilibrium, and the population totals 4000 individuals. (Let the frequency of • = q = 0.25, then use the Hardy-Weinberg genotype frequencies to calculate the predicted number of each genotype.) Use survival and reproductive rates that produces a slowly growing population and in which there is no selection. Do not use the same rates as in Exercise 5. If you wish, make such a run. You should be able to tell from the printout whether your initial population was in equilibrium. Do you understand the results you get?

	• •	• ◊	◊ ◊
Title	_____		
Carrying capacity:	_____		
Post-crash pop. size:	_____		
Initial population:	_____	_____	_____
Survival rates:	_____	_____	_____
Reproductive rates:	_____	_____	_____
Number immigrating:	_____	_____	_____
Percent emigrating:	_____	_____	_____

In all future runs of EVOLVE you should use initial populations that are in Hardy-Weinberg equilibrium — why?

Natural Selection

Natural selection is perhaps the most important of the evolutionary forces because it is the one most likely to lead to adaptation. Hence, we will devote the majority of the remainder of this chapter to an examination of patterns of selection.

Exercise 7. What Effect Does Increasing the Strength of Selection Have on the Evolution of an Advantageous, Dominant Allele?

Prediction: Setting up two populations, one with large and one with small differences between survival and/or reproductive rates of the two phenotypes, should show that evolution (that is, change in allele and genotype frequencies) proceeds more rapidly when the differences are larger.

Experiment: Use the following data for your two runs of EVOLVE:

1st run:

	Title: 7A. Sel for Dominant Allele (◊)		
Number of generations:	100		
Carrying capacity:	9999		
Post-crash pop. size:	7000		
	• •	• ◊	◊ ◊
Initial population:	1711	278	11
Survival rates:	20	24	24
Reproductive rates:	6	6	6

2nd run:

	Title: 7A. Sel for Dominant Allele (◊)		
Number of generations:	100		
Carrying capacity:	9999		
Post-crash pop. size:	7000		
	• •	• ◊	◊ ◊
Initial population:	1711	278	11
Survival rates:	20	34	34
Reproductive rates:	6	6	6

Results: Fill in the following tables from the data table produced by EVOLVE. Again, note that these tables are provided to help you learn what to look for and how to read EVOLVE's graphs. You will soon be able to see most of the clues on the screen without transcribing the data onto paper. If your are experienced with spreadsheets, you can also use Copy Window Data to copy the data to the Mac's clipboard. Then Paste the results into a spreadsheet and generate the tables below with that program.

Enter data on rates of change of allele and genotype frequencies:

Run with strong selection:

	Change in frequencies				
	Genotype frequency				Allele frequency
Generation	••	•◊	◊◊	•	◊ _____
1 -> 5	_____	_____	_____	_____	_____
6 -> 10	_____	_____	_____	_____	_____
11 -> 15	_____	_____	_____	_____	
16 -> 20	_____	_____	_____	_____	
21 -> 25	_____	_____	_____	_____	
31 -> 35	_____	_____	_____	_____	
51 -> 55	_____	_____	_____	_____	
71 -> 75	_____	_____	_____	_____	

Run with weak selection:

	Change in frequencies				
	Genotype frequency				Allele frequency
Generation	••	•◊	◊◊	•	◊ _____
1 -> 5	_____	_____	_____	_____	_____
6 -> 10	_____	_____	_____	_____	_____
11 -> 15	_____	_____	_____	_____	
16 -> 20	_____	_____	_____	_____	
21 -> 25	_____	_____	_____	_____	
31 -> 35	_____	_____	_____	_____	
51 -> 55	_____	_____	_____	_____	
71 -> 75	_____	_____	_____	_____	

a. Describe the pattern of change of allele frequency in the:

   Strong selection run:

   Weak selection run:

b. Describe the pattern of change of genotype frequencies in the

   Strong selection run:

   Weak selection run:

c. Fill in the following table of data on population size changes using EVOLVE's data
table:

	No. generations between crashes	
	Strong selection	Weak selection
Generation 1 -> 1st crash	_____	_____
1st crash -> 2nd crash	_____	_____
2nd crash -> 3rd crash	_____	_____
3rd crash -> 4th crash	_____	_____
4th crash -> 5th crash	_____	_____
5th crash -> 6th crash	_____	_____

Describe the pattern of population size changes in the:

   Strong selection run:

   Weak selection run:

d. Which run had the fastest rate of population growth at the end of the run?

   Run with strong selection   Run with weak selection

e. Describe the differences between runs with respect to:

   Allele frequency changes:

Population size changes:

Conclusions: Finally, summarize the results. That is, compare the evolution of dominant, advantageous alleles under strong and weak selection pressure.

Exercise 8. Does the Evolution (the Change of Allele and Genotype Frequencies) of an Advantageous, Dominant Allele Proceed More Rapidly Than That of an Advantageous, Recessive Allele with Comparable Survival and Reproductive Rates?

Prediction: (enter your own):

Experiment: You may use the input data and results from Experiment 7A as the control for the experiment that follows. However, if you wish, you may repeat 7A — you will gain more experience with the effects of randomness.

Use the following data; they are the same as for 7A, but the title has been changed and the advantageous allele (◊) is recessive:

	Title: 8A. Sel for Recessive Allele (◊)		
Number of generations:	100		
Carrying capacity:	9999		
Post-crash pop. size:	7000		
	• •	• ◊	◊ ◊
Initial population:	1711	278	11
Survival rates:	20	24	24
Reproductive rates:	6	6	6

Results: For both runs, compare the following (be sure to make specific comparisons):

Allele frequency changes:

Genotype frequency changes:

Population size changes:

Conclusion: Which type of allele evolved faster? Explain. Do you think the question, as it was initially phrased, was a useful one?

Exercise 9. How Does the Evolution of Incompletely Dominant Alleles Differ from the Evolution of Completely Dominant Alleles?

Not all alleles display dominance and recessiveness; perhaps the majority show more complex interactions. In this exercise you will look at another pattern of inheritance. We will define "incomplete dominance" as a situation where the heterozygote is exactly intermediate between the two homozygous phenotypes.

Prediction: (enter your own):

Experiment: You will use the results from Experiments 7A and 8A as the controls for this experiment, so change the title and make the survival rate of the heterozygote 22% compared to 20% and 24% for the homozygotes; leave all other variables the same.

Results: Compare the pattern of changes of allele and genotype frequencies, and of population size, of this experiment with those of 7A and 8A.

Conclusion: Summarize and explain these results.

Exercise 10. What Is the Evolutionary Fate of a Population in Which the Heterozygote Is the Most Fit Genotype (Heterosis)?

This question is a complex one and a complete exploration is beyond the scope of these exercises. If you go on to more advanced exercises you will have a chance to study this issue in more detail; here we will confine ourselves to a brief exploration with three experiments and then try to derive a general, qualitative conclusion.

a. Use the following data for your first run:

Title: 10A. Sel for Heterozygotes

	• •	• ◊	◊ ◊
Number of generations:	100		
Carrying capacity:	9999		
Post-crash pop. size:	7000		
Initial population:	1711	278	11
Survival rates:	20	24	20
Reproductive rates:	6	6	6

Prediction: (enter your own):

Results: record your results in the table that follows 10C.

b. Now, try an experiment with these data (keep other variables the same as above):

Title   10B. Sel for Heterozygotes

	• •	• ◊	◊ ◊
Initial population:	11	278	1711

Prediction: (enter your own):

c. Now, try an experiment in which one homozygote is sterile:

Title: 10C. Sel for Heterozygotes

Number of generations:	100		
Carrying capacity:	9999		
Post-crash pop. size:	7000		
	• •	• ◊	◊ ◊
Initial population:	1711	278	11
Survival rates:	20	24	20
Reproductive rates:	6	6	0

Prediction: (enter your own):

Results: What were the final allele and genotype frequencies?

	Genotype frequency			Allele frequency	
Experiment	••	•◊	◊◊	•	◊
10A	_____	_____	_____	_____	_____
10B	_____	_____	_____	_____	_____
10C	_____	_____	_____	_____	_____

Conclusion: Summarize your three experiments; what is the general fate of alleles when the heterozygote is most fit? Explain.

Genetic Drift

Genetic drift is defined as random changes of allele or genotype frequencies caused by "sampling error" in populations of finite size. In essence, the processes of survival and reproduction take some of the alleles and genotypes of a population to make up the next generation. Even in a population that is not subject to selection or any other evolutionary force, the laws of chance dictate that the frequencies in one generation will not be exactly the same as those of the previous or succeeding generations. If the population is small, then chance will play a bigger role. The same principle applies when you flip coins: if you tossed a coin 4 times, you wouldn't be surprised to get 3 heads and 1 tail, but if you tossed it 100 times you would be suspicious if you got 75 heads and 25 tails.

Randomness is a difficult thing for us to grasp; if you tried to say 100 random digits, statistical tests would show that your numbers were biased in some way. Similarly, looking at graphs of allele frequencies, it is difficult to tell if there are random changes. The next experiments should help you get a better feel for randomness.

Exercise 11. What Effects Does Population Size Have on Allele Frequencies?

Experiments: Make at least two runs of EVOLVE to compare changes of allele frequencies in populations of different sizes. To limit population sizes, you should input appropriate values for maximum and post-crash population sizes. For example, runs with K and post-k set to 20 and 10, or 500 and 250, or 2000 and 1000, or 5000 and 2500 would permit comparison of populations of very different sizes, yet all would suffer crashes of 50%. Initial populations should be in Hardy-Weinberg equilibrium, have initial allele frequencies of 0.50, and be equal to the post-crash population size. All genotypes should, of course, have equal survival and reproductive rates — we do not want selection to confuse our interpretation of the effects of population size.

Results: Describe the pattern of changes of allele and genotype frequencies shown by your experiments.

If you did not know the data that had been used to start the experiments, would you think that any of the graphs showed significant changes in allele or genotype frequencies? Use your fingers to display 10- or 20-generation segments of your graphs and try to decide whether evolution occurred during that time.

Was there any pattern to which allele or genotypes increased in frequency?

Conclusions: What impact does population size have on evolution?

Gene Flow

Gene flow is the net movement of alleles from one population to another. Because of uncertainty over the actual amount of gene flow in nature, there is considerable debate over its actual importance. Regardless of what actually happens, EVOLVE can illustrate what might happen if gene flow were significant.

Exercise 12. What Is the Effect of Gene Flow on Evolution?

In this exercise we will conceive of our population as living on an island or mountain top surrounded by an inhospitable area. Individuals which disperse away from our population may be regarded as having died (that is, having been removed by selection); although they may survive and contribute to another population, they are lost to the one we are studying.

Let us suppose our population initially contained only the ◊ allele, and began receiving immigrants from another population that contained a high frequency of • alleles. If the • allele produced a phenotype which dispersed readily and had no effect on survival or reproduction, then the gene flow would favor that allele (that is, cause it to increase in the population).

Set up you EVOLVE data as follows:

	• •	• ◊	◊ ◊
Title: 12. Gene Flow			
Number of generations:	100		
Carrying capacity:	9999		
Post-crash pop. size:	7000		
Initial population:	0	0	8000
Survival rates:	20	20	20
Reproductive rates:	6	6	6
Number immigrating:	1000	100	0
Percent emigrating:	0	0	0

Prediction: (enter your own):

Results: Describe the changes in allele and genotype frequency.

Conclusion: What is the effect of gene flow on allele and genotype frequencies? In this experiment we neglected the fact that •-containing individuals would probably

emigrate at higher rates than ◊◊ individuals. How would the results of this experiment differ if the emigration rates were set to 30%, 30%, and 0% for the ••, •◊, and ◊◊ genotypes, respectively?

Mutation

Mutation is of great interest to biologists because it is the ultimate source of the genetic variability that is the raw material of evolution. As you might expect, the study of mutation is a complex one and for many reasons EVOLVE cannot be used to examine such things as mutation rates and mutational equilibrium. However, it does allow you to study the fate of a mutation that has already occurred, and that is the objective of this exercise.

Exercise 13. What Is the Fate of Advantageous Mutant Alleles?

You have already made runs with advantageous dominant and advantageous recessive alleles, using initial advantageous allele frequencies of 5%. Using the results of those runs as controls, compare the fate of an advantageous recessive mutant or that of an advantageous dominant mutant and explain what happens.

Prediction: (enter your own):

Experiment: To simulate a population in which a mutation has just occurred, you will want an initial population with many ••s, 1 •◊, and 0 ◊◊s. The phenotypes of the dominant and recessive advantageous alleles should be the same as those used in Exercise 7 or 8 (whichever will be your control). You may wish to make additional runs.

Results: Summarize the results of your runs.

Conclusion: What is the typical fate of a new, advantageous, dominant mutant? Most mutations that affect the phenotype are believed to be recessive and deleterious. What then is the fate of most such mutations? What about mutations that have no effect on the phenotype?

Combining Evolutionary Forces

The exercises up to this point have attempted to help you gain a feel for how individual evolutionary forces operate by themselves. However, no real population is subject to only one force at a time. In the following exercises you will compare the effects of previous experiments with ones in which more than one force is operating.

Exercise 14. Drift and Selection

The question of the relative importance of genetic drift and natural selection is a hot topic among evolutionary biologists these days. Here you will get a chance to see what the fuss is all about.

Set up an experiment which uses survival and reproductive rates from one of the experiments in Exercises 7-10, and the smallest K and post-k values you used in Exercise 11. Compare the results with each of the first two (which are now the controls for Exercise 14).

Prediction: (enter your own):

Results: Describe the overall changes in allele frequency and the generation-to-generation fluctuations. Compare them with the allele frequency changes in the two earlier exercises. Compare your results with those of your classmates.

Conclusions: What is the effect of drift on selection? Can an advantageous allele be lost as a result of drift? Can a deleterious allele increase in frequency?

Exercise 15. Selection and Gene Flow

One of the major controversies in evolution is why differing selection regimes in different environments do not cause widely spaced populations of one species to diverge from each other.

The Osprey, for example, is a predatory bird that feeds on fish and is found around seashores, lakes, and rivers around the world in temperate and tropical regions. One would expect that natural selection over such a range of habitats must select for different characteristics, yet Ospreys look the same all over the world. One possibility, of course, is that natural selection acts the same way in all Osprey populations. Many evolutionists, however, have felt that gene flow can act to hold populations together even though natural selection would tend to cause them to become different. Other evolutionists believe that there is not enough gene flow to counteract selection and that there must be other reasons for the similarity of species over wide geographic areas. Here is an experiment that addresses this controversy.

Experiment: Consider a population on an island where selection favors a recessive allele, ◊. Perhaps individuals carrying the dominant • allele have seeds with long hairs like milkweed that tend to be blown off the island, while ◊◊ individuals produce seeds with short hairs that fall directly to the ground. The population in Experiment 8A illustrates what would happen to the population if it were isolated — the • alleles would decrease and the population would come to consist of plants with short-haired seeds. To see the effect of gene flow, set up another set of experiment like the following in which hairy seeds are blown onto the island from a mainland where selection favors mobile seeds:

Title: 15A. Sel vs Gene Flow (0.5%)			
Number of generations:	100		
Carrying capacity:	9999		
Post-crash pop. size:	7000		
	• •	• ◊	◊ ◊
Initial population:	1711	278	11
Survival rates:	20	24	24
Reproductive rates:	5	6	6
Number immigrating:	16	4	0
Percent emigrating:	0	0	0

Title: 15B. Sel vs Gene Flow (5%)			
Number immigrating:	160	40	0

Title: 15C. Sel vs Gene Flow (50%)			
Number immigrating:	1600	400	0

Prediction: (enter your own):

Results: Describe the changes in allele and genotype frequencies in these populations and compare them with the changes in Experiment 8A.

Conclusion: Does this experiment support the hypothesis that gene flow could reduce the divergence of populations?

Chapter 6. Intermediate Exercises

This chapter provides some additional exercises to sharpen your understanding of evolution. These exercises differ from earlier ones by being more quantitative and more abstractly formulated. The impact of selection on population sizes will also be examined in this chapter. We assume a fair amount of sophistication on the part of the student and do not set up experiments in as much detail as in the previous chapter.

Because evolution is the result of the interaction of many statistical phenomenon, and because EVOLVE simulates random processes, one or even several runs with a particular set of variable values may not show all of the variation needed to understand the phenomena being studied. Many of these exercises are best done by teams of students who divide up a number of runs and then pool their results.

Preliminary Exercises

Like those in Chapter 5, these initial exercises, begin with a few questions designed to test your understanding of some basic concepts underlying EVOLVE. In this case we will be using the more abstract concepts of absolute and relative fitness, studying their impact on growth of population size. Use the following tables to answer Questions 1-5.

a.

Phenotype	••	•◊	◊◊
Surviv. rate	25.0%	25.0%	20.0%
Reprod. rate	3.0	3.0	3.0

Genotype

b.

Phenotype	••	•◊	◊◊
Surviv. rate	50.0%	30.0%	30.0%
Reprod. rate	2.0	3.0	3.0

Genotype

c.

Phenotype	••	•◊	◊◊
Surviv. rate	40.0%	40.0%	58.0%
Reprod. rate	1.0	2.0	3.0

Genotype

d.

Phenotype	••	•◊	◊◊
Surviv. rate	12.0%	18.0%	24.0%
Reprod. rate	9.0	9.0	9.0

Genotype

e.

Phenotype	••	•◊	◊◊
Surviv. rate	60.0%	75.0%	65.0%
Reprod. rate	2.0	1.0	1.0

Genotype

f.   None of above

1. Calculate the absolute and relative fitnesses of each genotype in each table:

a. Genotype	Fitness Absolute	Relative	b. Fitness Absolute	Relative
• •	_____	_____	_____	_____
• ◊	_____	_____	_____	_____
◊ ◊	_____	_____	_____	_____

c.

Genotype	Fitness Absolute	Relative	d. Fitness Absolute	Relative
• •	_____	_____	_____	_____
• ◊	_____	_____	_____	_____
◊ ◊	_____	_____	_____	_____

e.

Genotype	Fitness Absolute	Relative	f. Fitness
• •	_____	_____	
• ◊	_____	_____	
◊ ◊	_____	_____	

2. Which of the tables in Question 1 shows dominance with respect to relative fitness?

a.   b.   c.   d.   e.   f.

3. Which of the tables in Question 1 shows heterosis for relative fitness?

a.   b.   c.   d.   e.   f.

4. Which of the tables in Question 1 shows incomplete dominance for relative fitness?

a.   b.   c.   d.   e.   f.

5. Which of the populations in Question 1 would become extinct if there were 4000 individuals, the frequencies of both alleles were equal to 0.5, and the genotypes were initially in Hardy-Weinberg equilibrium?

a.   b.   c.   d.   e.   f.

Exercise 16. Selection via Reproduction vs Selection via Survival.

In our formulation of the concepts of relative and absolute fitness, we multiply reproductive rate by survival rate to obtain single coefficients of fitness for each allele and genotype. We then compare and standardize those coefficients by dividing by the largest. While this approach is a useful and powerful one, it is always worth asking if the technique used to study some phenomenon might obscure interesting points.

One such question is whether or not it makes a difference if a given pattern of selection is achieved by survival rate or by reproductive rate, or by a combination of both. For example, the absolute fitness of a genotype with 20% survival and an average of 6

offspring = .2 x 6 = 1.2. The same fitness could be achieved by 40% survival and 3 young, or 30% survival and 4 young, or 60% survival and 2 young.

Would evolution proceed the same in two populations with identical fitness and selection coefficients, if one achieved that pattern of selection with high survival and low reproduction, and the other achieved it with low survival and high reproductive rates? Write down your predictions of what you expect to happen.

Set up and run a series of EVOLVE experiments to examine this question. Be sure to examine graphs of genotype numbers and population size.

What do your results suggest about the usefulness of fitness and selection coefficients in studying the evolutionary biology of real organisms?

Ecologists studying **life-history patterns** (groups of traits related to longevity and reproduction) have found correlations of traits that make up alternative life-histories. For example, a short life-span is often correlated with rapid maturation, production of large numbers of small young with little parental care, and a single episode of reproduction. Species exhibiting such traits are often "weedy" colonizers of unstable habitats and are sometimes called **r-selected.** Opposite traits (long life-span, intensive parental care of a few, large young, repeated reproductive attempts) are associated with **K-selected** species, which often live in stable habitats. How do your results bear on the ecological concepts of r- and K-selected populations? Look up these concepts in your textbook if you need more background on these concepts.

Exercise 17. Exploring Heterosis.

In Exercise 10 you briefly examined the pattern of evolution of a population with heterosis. However, that did not really reveal the wealth of interesting phenomena inherent in this pattern of inheritance. This exercise is designed to give you a framework for exploring the phenomena of heterosis.

The basic question is: What is the evolutionary fate of a population showing heterosis? Compare the changes in allele frequencies in a populations where the heterozygous •◊ genotype is more fit than either of the homozygotes.

A convenient metaphor that may be helpful in planning experiments is to think of yourself as exploring a landscape on the plane illustrated in Figure 6-1 below. The horizontal axis represents the absolute fitnesses of the two homozygotes. On the far left, the ◊◊ homozygote's fitness is zero and the •• has very much higher fitness. In the middle, the fitness of the two homozygotes is equal. On the right, the •• has a fitness of zero and the ◊◊ is higher (but still less than that of the heterozygote).

Four runs are outlined for you on the figure and in the table below. In Run 1 and Run 2, the homozygotes have equal fitnesses of 1.5; the only difference between the two is that Run 1 starts with an initial • frequency of 5%, while Run 2 starts with a • frequency of

95%. Run 3 starts with the same low allele frequency as Run 1, but the fitness of the ◊ ◊ homozygote is zero. In Run 4, the initial allele frequency is the same as Run 3, but the •• genotype has a lower (but non-zero) fitness and the ◊◊ genotype has a higher fitness. You should make these runs and other runs of your own design to fill in the unexplored open areas of the figure.

Figure 6-1. A Metaphorical Landscape for Exploring Heterosis.

Make several runs in which the fitnesses of • and ◊ homozygotes are equal, but which start with different initial allele frequencies (for example, Runs 1 and 2 on the table below and their corresponding numbers in the graph above). Next, make runs with the absolute fitnesses indicated in Runs 3 and 4 on the next page and indicated by the "3" and "4" on the graph. Note the final allele frequencies of the populations and mentally summarize the results. Make a prediction about the final allele frequencies at some other point on the plane and test your prediction with another run of EVOLVE. In this way you can explore all of the regions on the surface before writing your answer to the question.

Run No.	Max pop.	Post-crash	Initial population			Initial frequency of • allele
			• •	• ◊	◊ ◊	
1	____	____	____	____	____	0.05

Survival rates			Absolute fitness			Frequency of • in generation 100
• •	• ◊	◊ ◊	• •	• ◊	◊ ◊	
50	50	50	1.5	2.0	1.5	____

Reproductive rates			Relative fitness		
• •	• ◊	◊ ◊	• •	• ◊	◊ ◊
3	4	3	____	____	____

	Initial population	Initial frequency

Run No.	Max pop.	Post-crash	Initial population			of • allele
			• •	• ◊	◊ ◊	
<u>2</u>	_____	_____	_____	_____	_____	<u>0.95</u>

Survival rates			Absolute fitness			Frequency of • in generation 100
• •	• ◊	◊ ◊	• •	• ◊	◊ ◊	
50	50	50	1.5	2.0	1.5	_____

Reproductive rates			Relative fitness			
• •	• ◊	◊ ◊	• •	• ◊	◊ ◊	
_____	_____	_____	_____	_____	_____	

Run No.	Max pop.	Post-crash	Initial population			Initial frequency of • allele
			• •	• ◊	◊ ◊	
<u>3</u>	_____	_____	_____	_____	_____	<u>0.05</u>

Survival rates			Absolute fitness			Frequency of • in generation 100
• •	• ◊	◊ ◊	• •	• ◊	◊ ◊	
_____	_____	_____	1.5	2.0	0.0	_____

Reproductive rates			Relative fitness			
• •	• ◊	◊ ◊	• •	• ◊	◊ ◊	
_____	_____	_____	_____	_____	_____	

Run No.	Max pop.	Post-crash	Initial population • •	• ◊	◊ ◊	Initial frequency of • allele
<u>4</u>	____	____	____	____	____	<u>0.95</u>

Survival rates • •	• ◊	◊ ◊	Absolute fitness • •	• ◊	◊ ◊	Frequency of • in generation 100
____	____	____	0.5	2.0	1.75	____

Reproductive rates • •	• ◊	◊ ◊	Relative fitness • •	• ◊	◊ ◊
____	____	____	____	____	____

Exercise 18. Modeling the Real World — Sickle-Cell Anemia

It is always a good idea to test your scientific models against the real world. Often data needed to do such a test are available in the scientific literature. In Chapter 3 of this manual, we used sickle-cell anemia as an example of a somewhat complex evolutionary situation, but we did not fill in our tables with actual data on survival and reproduction.

Do some library research and see if you can find "hard" data on the survival and reproductive rates of people with each of the sickle-cell genotypes in areas with and without endemic malaria. Use those data as inputs to EVOLVE, and see if the program predicts the actual allele frequencies observed in those populations. Comment on the closeness of the results and discuss factors that might have affected EVOLVE's accuracy.

Exercise 19. Plotting Δq vs q.

As you may know from an examination of texts on theoretical population genetics, mathematical models of selection frequently derive the Δq, or change in allele frequency per generation, as a function of allele frequency. Even though you may not have studied such derivations, it is interesting to tabulate and graph these statistics. Examination and discussion of sample curves for EVOLVE experiments may help you understand the reasons for the shape of allele frequency curves over time.

We suggest you examine the allele frequency graphs of comparable experiments, for example, 7A and 8A, which involved selection for dominant and recessive alleles,

respectively. You can use Copy Window Data to copy the summary data to the Mac's clipboard, then paste them into a spreadsheet or statistics program. Tabulate the allele frequency (q) of the allele that increases in frequency. Subtract adjacent generations' values of q to obtain Δq; then do a scatter plot of Δq vs q. Be sure to note the survival and reproductive rates of the experiment. If members of a class each do a separate experiment, you will have abundant material for discussion.

For example, ask yourself what values of q have the highest values of Δq in Experiments 7A and 8A. Can you see how the graphs of allele frequency vs time that EVOLVE produced relate to the plots of Δq vs q? Do the latter show you why the steep part of the allele frequency vs time graphs occur early and late in the experiment? Compare the results from 7A with a similar graph of 7B. How does increasing the strength of selection affect Δq?

Comparison of other Δq vs q graphs is also revealing. You might find it especially interesting to look at 10A, 10B, and 10C.

Exercise 20. Plotting Δq vs Population Size.

When you are studying drift, it is more useful to graph Δq vs population size. You should take a sample of 10 or so pairs of years from the results of each experiment you did for Exercise 11, tabulate population size and allele frequency, then do a scatterplot of allele frequency vs population size. You could also take the data from Exercise 4 during the period when the population was expanding — it will fill in the space between the points from Exercise 11. (Give some thought to whether you should plot the size of the population in the first or last year, or if you should take the average of the two).

Exercise 21. Examining q at a Time t for a Large Number of Populations of the Same Size.

Another way to look at drift is to do a series of experiments like those in Exercise 11 but with various random numbers. Pooling class data is especially useful here. Record the allele frequency of one allele every 25 generations. When you have results from 10 or more experiments of the same size, make histograms of allele frequencies — you see graphically that genetic drift can indeed result in significant changes in allele frequency.

Chapter 7. Advanced Exercises

The exercises in this chapter are designed to help you develop a more complete understanding of the basic design of EVOLVE, of the difficulties of determining what types of selection might be occurring in nature, and of how one might study evolution statistically.

Exercise 22. The Model Underlying EVOLVE

As noted in Chapter 1, theoretically minded evolutionists have continued to try to develop models of evolutionary and ecological processes despite (or because of) difficulties with experimentation. Their models are conceptual, rather than experimental, often mathematically sophisticated, and a complete understanding of some of them requires the use of probability, calculus, matrix algebra, Markov Chains, and/or game theory. Such models clearly are beyond the scope of many biology courses. Nevertheless, modeling is too useful a process to ignore.

By the time you get to this exercise you should have spent quite a bit of time using EVOLVE, but you probably do not really understand how it can produce realistic results in such a wide variety of evolutionary situations. What we shall do in this exercise is establish a simple, intuitive model and explore the results it produces in one evolutionary situation. This will give you an understanding of one approach to modeling. An understanding of the basic approach of this type of model is very important in more detailed analyses of its results.

Before going to class, consider how you would model the evolutionary situation described below. Discuss this assignment with others before class if you wish. You are not expected to produce a finished product, but should try to map out how you would approach the problem. Take your notes to class and be prepared to discuss your approach. In class you can set up a model and run it by hand. After getting results we will see how this model relates to EVOLVE. Please take a calculator to class.

Restrict your thinking to a species which is hermaphroditic (has both male and female organs) and diploid, reproduces sexually, and must mate with another individual.

Consider a life cycle where individuals are born during a restricted breeding season, mature for one year, mate, produce young, and die. Consider only changes in the frequency of two alleles (• and ◊) at one locus. If it helps you visualize an abstract problem, think of the ◊ alleles producing spines which protect heterozygotes and homozygotes against predation. However, in homozygotes the ◊ alleles produce a shorter copulatory organ; along with the spines, this reduces effectiveness of copulation. Consequently, ◊◊ individuals have a higher survival rate than •• and •◊ individuals, but a lower reproductive rate. Heterozygous individuals (•◊) have the best

of both effects — protection and a high reproductive rate. Start with a population with the following composition and average survival and reproductive phenotypes:

Genotype	Number	Mean survival rate	Mean reproductive rate
• •	245	0.2	5.1
• ◊	210	0.3	4.6
◊ ◊	45	0.3	4.1
Total	500		

You need not try to formulate one equation for such a population; you may find it easier to consider the basic population process of birth, survival, mating, and death.

Questions: (Bring brief, specific answers to class for discussion):

1. What do you predict will happen to the population size over time?

2. What do you predict will happen to the frequencies of the two alleles?

3. What do you predict will happen to the frequencies of the three genotypes?

4. How would you set out to convince a skeptic that your prediction is correct?

Your instructor will help you discuss your approaches and come up with an approach to model this evolutionary scenario. You can then calculate the results to see how your answers compare to the model the class develops.

Exercise 23. Statistical Comparisons of EVOLVE's Results with Theory

One of the major difficulties scientists have is to evaluate whether data that they have gathered in experiments "fit" the predictions of theory. Students with some background in statistics may find that EVOLVE provides an opportunity to gain experience with statistical tests of "goodness of fit." If you have not already done so, you may want to see the BioQUEST statistics module for some approaches.

Pick an evolutionary situation about which you can make mathematical predictions and compute values of some parameter. Then do one or more EVOLVE experiments that provide you with data about that situation. Graph the predicted values against those provided by EVOLVE and evaluate the "goodness-of-fit." For example, when selection operates against a lethal or sterile recessive (that is, the selection coefficient = 1 and relative fitness = 0), then the formula for change in gene frequency becomes:

$$\Delta q = -q^2 / (1+q)$$

You might calculate the values of Δq at intervals of q between 0 and 1 according to this equation, then plot the results from one or more EVOLVE experiments. What statistical tests would let you test the hypothesis that the EVOLVE data resulted from a population subjected to selection against a lethal (or sterile) recessive? Do your results refute the hypothesis?

Exercise 24. Inferring Pattern of Selection from "Field" Data

Another lesson about the difficulty of studying evolution can be learned from examining EVOLVE output without seeing the input data. If your instructor gives you printouts of an EVOLVE experiment, either generation summaries or graphs, could you tell what the input values were? It is instructive to break your class or lab up into several teams and for each team to solve the puzzle. You might test your ideas by making your own EVOLVE experiments to try to duplicate the result you were given. Would your task be easier or harder if the printout were from a trial with a small population?

Do all teams come up with the same ideas? Going one step further, what does your experience with this exercise (and other EVOLVE exercises, especially 22) tell you about scientists trying to explain real phenomena in nature?

Remember that in the "real world" we rarely have such good data on allele or genotype frequencies as EVOLVE provides.

Chapter 8. Program Notes and Setting up Experiments

Introduction

This chapter contains reference material on each of the parameters you may input to EVOLVE, as well as notes on output and the help system. The emphasis is on how each parameter relates to evolution, and on practical aspects of setting up experiments. There is also some discussion of how to interpret the various output graphs. As with most reference works, you will probably not find it useful to read through this chapter trying to absorb everything in it. Rather, skim through it to get an idea of its contents, then refer back to it when you have a particular question.

Getting Help with EVOLVE

There are three ways that the user can get help with the EVOLVE program.

1. Information about inactive (gray) menu items: If you do not understand why a particular menu item is not available, you can click on an inactive item. A small window will appear giving information on the item and indicating why it cannot be chosen at this time. For example, choosing the Cut item in the **Edit** menu displays the following message: *"This option is only active when some text is selected."*

2. General information: For help with a particular menu item or about a window on the screen, you can enter Help mode. To enter Help mode, hold down the Command (or) key while typing a question mark. (If you have a help key on your keyboard this will have the same effect.) The cursor will change to a question mark. Clicking on any item in a menu (all items are enabled in this mode) gives general information about that item, including what would happen if that item were chosen. For example, in Help mode, choosing the Cut item displays the following message: *"This will delete the selected material from a notepad and retain it in the clipboard so that it can later be pasted into another notepad or into a word processor."*

Similarly, clicking anywhere in the front-most window while in Help mode will display some general information about that window. This information includes a description of the contents of the window and explains the function of any buttons or other controls on the window. For example, clicking on the Parameter History Window while in Help mode displays the following message: *"This is a scrollable list of all the parameters used in the experiment so far. This is not editable but you can make an editable copy of it by choosing the 'Copy Window Data' option on the 'Edit' menu."*

Help mode will automatically turn off once you have clicked on an item.

3. To browse through all of the information available about EVOLVE: Choose Help with Evolve under the menu on the far left. This allows you to choose from a list of topics and display information on the topic that interests you. Buttons at the bottom of the window let you move to additional topics in the list.

Program Input

Title

Be sure to enter a brief, descriptive title for your experiments. You will eventually accumulate many different experiments and the titles will help you keep them straight. For example, if a number of people are storing their results on one disk, you should put your initials in the title so you won't get your results mixed up with someone else's.

Seed

This is a whole number between 0 and 32767. It is used by the computer to simulate random mating, the effects of weather, and other "random" factors.

You may let EVOLVE pick the seed for you, or you can type a seed in if you wish. 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. A thorough understanding of the seed is not necessary for you to use EVOLVE, you may find the following background helpful.

EVOLVE takes the "seed" you enter in the Seed box and performs a series of mathematical operations on it to produce a second "random" number. This number is used by EVOLVE in various ways — for example, to randomly pick pairs of parents from the adult population, and to let the survival rate vary slightly from the average values you entered in the Change Parameters Window. This second number is also used as the seed for the third "random" number, and so on.

These numbers are not really random, because the mathematical operations will always produce the same sequence of numbers if the same seed is used — the sequence of numbers is determined by the value of the initial seed. However, if you examine a listing of the numbers produced by such a "random number generator" you would not see a relationship between them, so in this sense they are random.

This may become clearer if you consider another, more commonly used, random number generator, a tossed coin. If you knew enough physics, and had enough information on the initial position, size, shape, and mass of the coin, along with the rotational velocity, angle, and velocity of the toss, and such things as wind direction

and velocity, you could predict heads or tails before the coin landed. In this sense, the result of a coin toss is determined, yet for all practical purposes we can regard the outcome as random.

Again, consider how EVOLVE uses random numbers. When the program reduces the number of young to obtain the number of adults, it takes the survival rate and seed you gave it and produces an actual "survival rate" that is multiplied by the number of young to obtain the number of adults. The actual survival rates used by EVOLVE over a number of generations will have an average value very close to the value you gave it, although the value of a given generation will be uncertain. If you input 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 may seem complicated, but it is a practical way of simulating the sort of variations that occur in nature. For example, weather might be mild for a couple of years, then harsh, or predators that take advantage of one phenotype might be unusually abundant in one generation and unusually scarce in another. If climate is what you expect, weather is what you get; you enter "climate" into EVOLVE's menus, the random number generator gives you "weather."

Number of Generations

EVOLVE will continue an experiment for the number of generations specified by this number. Although you may stop an experiment at any time, it may be useful to have EVOLVE stop at a predetermined time. Because it takes longer to run many generations, you may wish to explore combinations of data by running for, say, 50 generations instead of the full 200. Also, there are times when you may want to change variable values on a specific schedule.

Starting Population

These three numbers determine the composition and size of the starting adult population of generation 1. They must not total more than 8000.

For example, suppose you wish to study the fate of a new mutant allele (◊) in a population of 1000 •• individuals. You should use 999, 1, and 0 as the initial numbers of ••, •◊, and ◊◊ genotypes, respectively. A new mutation obviously would first exist in heterozygous condition.

If you wish to establish a population of 2000 with the frequency of the • allele equal to 0.40, use 320, 960, and 720. These numbers can be obtained using the Hardy-Weinberg proportions — you should be able to determine them. (See "Calculation of Fitness and Selection Coefficients" in Chapter 9, or your textbook, if you do not know how to do this.)

Survival and Reproductive Rates: Selection, Pattern of Inheritance, and Population Growth Rates

Survival rates determine the average number of young of a particular genotype that survive to adulthood. They may vary from 0% for a lethal genotype to 100 for a genotype in which 100% of the young survive to adulthood. The former is not unusual in nature, but the latter is very unlikely.

Reproductive rates determine the average number of young per individual. They may range from 0 offspring per parent for a sterile genotype to 10 offspring per parent for a fertile one. The limit of 10 young is arbitrary, imposed to make EVOLVE simpler and let it run more quickly.

These are very important variables which determine the pattern of inheritance, and of natural selection, and greatly influence rate of population growth. To understand population genetics and EVOLVE, you absolutely *must* understand them thoroughly.

Pattern of Inheritance — Choice of appropriate values for reproduction and survival will allow you to simulate any pattern of inheritance possible for a single locus with 2 alleles (for example, dominance and recessiveness, codominance, or heterozygote superiority). One definition of a dominant allele is that it is one which produces the same phenotype when heterozygous as when homozygous.

For example, if the survival rates were 20%, 20%, and 50% for ••, •◊, and ◊◊ genotypes respectively, then the • allele would be dominant and the ◊ allele would be recessive. Note that dominance has nothing to do with which genotypes have the highest survival or reproductive rates, only with the homozygote that the heterozygote resembles. In this case, the heterozygote has the same traits as the homozygous ••, so the • allele is dominant. To take another example, reproductive rates of 5, 4, and 3 would simulate a pattern of inheritance where heterozygotes are intermediate between the homozygotes, that is, incomplete dominance. The following table provides more examples of the basic patterns of inheritance:

	Example	Genotype			Notes
		••	•◊	◊◊	
1.	Survival rate	30	30	20	• Dominant, advantageous;
	Reproductive rate	4	4	4	◊ Recessive, deleterious
2.	Survival rate	20	20	30	• Dominant, disadvantageous;
	Reproductive rate	4	4	4	◊ Recessive, advantageous
3.	Survival rate	30	30	20	• Dominant, advantageous;
	Reproductive rate	4	4	4	◊ Recessive, deleterious

4.   Survival rate   20   25   30   • Incompletely dominant, disadvantageous,

         Reproductive rate   4   4   4   ◊ Incompletely dominant, advantageous

5.   Survival rate   20   20   30   • Dominant, disadvantageous, ◊ Recessive, advantageous;

         Reproductive rate   4   5   5   ◊ Dominant, advantageous, • Recessive, deleterious

         Overall, the alleles demonstrate pleiotropy (multiple effects) and codominance (both have their effects in heterozygotes).

6.   Survival rate   20   40   30   • Dominant, disadvantageous, ◊ Recessive, advantageous;

         Reproductive rate   4   6   5   ◊ Dominant, advantageous, • recessive, deleterious;

         Overall, the alleles demonstrate overdominance with respect to fitness and are heterotic (show heterosis, or heterozygote superiority).

7.   Survival rate   20   10   30   • Overdominant, disadvantageous;

         Reproductive rate   4   6   5   ◊ Overdominant, advantageous;

         Overall, the alleles demonstrate overdominance with respect to fitness and show heterozygote inferiority.

Population Growth Rate — The relationship between reproductive rates and survival rates, what evolutionary biologists call "absolute fitness," determines the direction and rate of change of population size. You should choose survival and reproductive rates in such a way that your populations grows slowly. For example, you wouldn't normally consider an experiment where all genotypes have survival rates of 65% and reproductive rates of 1. For every 100 young born, there would be about 65 adults; each adult would produce an average of 1 young for the next generation (a total of 65 young). Given enough time, the population would decline to extinction.

In general, if the product of survival rates x reproductive rates is greater than 100, the population will not become extinct. Of course, in most runs the survival and reproductive rates will not be the same, so the exact pattern of population change will depend on the relative proportions of the various genotypes in the population as well as on the particular survival and reproductive rates.

Pattern of Selection — This also depends on the relative values of both survival and reproductive rates of each of the genotypes, what evolutionary biologists call "relative fitness." Suppose the survival rates were 60%, 40%, 40%, and the reproductive rates were 3, 2, 2 for the ••, •◊, and ◊◊ genotypes, respectively. The • allele would be recessive (because the heterozygote has the same characteristics as the ◊◊ homozygote),

but advantageous (because the •• homozygote has the highest survival and reproductive rates). Natural selection would favor the •• homozygotes.

As you work with EVOLVE you will gain a better feel for how these parameters interact. From the start, however, you may easily determine the *absolute fitness* of a genotype by multiplying the number of young by the survival rate. Products greater than 100 indicate that the genotype will tend to increase in number, while products less than 100 indicate that the genotype will tend to decrease in number. Thus, absolute fitness tells you how the actual *numbers* of individual organisms are likely to change. However, the absolute fitness of a genotype does not necessarily tell you whether the *proportion* of that genotype in the population will rise — that is, whether that genotype is favored by natural selection.

The relative sizes of the absolute fitnesses of the genotypes will indicate how the proportions of particular genotypes will tend to change over time. Dividing the absolute fitness of each genotype by the highest absolute fitness yields the *relative fitness.* The maximum value of the relative fitness will, of course, be one. Genotypes with relative fitnesses less than one will tend to decrease in frequency.

Take as an example the following:

	Genotype		
Phenotype	• •	• ◊	◊ ◊
Survival rate	20%	20%	14%
Reproductive rate	5	6	7
Absolute fitness	100%	120%	98%
Relative fitness	0.83	1.0	0.82

In this table, the survival and reproductive rates would be input to EVOLVE. The • allele would be dominant and advantageous with respect to survival, but incompletely dominant and deleterious with respect to reproduction. What would happen to the population? The ◊◊ genotype, although it has a low survival rate, has a high reproductive rate. Would the higher reproductive rate make up for the reduced survival? Would the ◊ allele increase or decrease in the population? Would it become extinct? Would the • allele become extinct? Or would both remain in the population?

Multiplying survival rates by reproductive rates gives us the third row — the absolute fitness. Here it is clear that the high reproductive capacity of the ◊◊ genotype does not completely compensate for the low survival, for the absolute fitness of 98 is less than 100. By dividing all of the absolute fitnesses by 120 (the highest) you can see that the •• homozygotes have only 83% of the fitness of the heterozygotes, while the ◊◊ individuals have only 82% relative fitness. By the time you finish your experiments with EVOLVE, you should be able to predict the outcome of such an evolutionary situation by calculating the relative fitnesses.

Gene Flow

To simulate gene flow, the movement of alleles from one population to another, you must enter rates of immigration (arrival) and emigration (departure) for each genotype. The number immigrating may vary from 0 to 4000 adults of each genotype per generation. The percent emigrating may vary from 0% for stay-at-homes, to 100% for genotypes with wanderlust.

Note that immigration is a constant and is independent of the density of the simulated deme, while emigration removes a constant proportion of each genotype in each generation. This may seem an unusual or odd situation, but it could be visualized in the following way: Suppose our model population is a plant similar to a dandelion, and the seeds have plumes for wind dispersal. The alleles might be conceived of as affecting the length of the plumes, and therefore the tendency to disperse (be blown away). Our model deme could be thought of as existing on a small island which is downwind of a large island or continent with a stable population. In such a situation, immigration could be viewed as constantly adding a particular number of individuals of each genotype, with long-plumed seeds arriving at a higher rate than short-plumed seeds. The short-plumed seeds would emigrate, that is, be blown away, from the island at a lower rate than the former.

Maximum Population and Post-Crash Population: Population Size

Along with rates of survival and reproduction (and gene flow, if any), the values you enter for maximum and post-crash population determine the size of the population. In EVOLVE, the maximum population may range from 10 to 5000, while post-crash size may range from 2 to 4000. Obviously, the maximum population must be larger than the post-crash size.

The maximum population is roughly equivalent to the ecologist's Carrying Capacity (the greatest number of individuals of a population that can be sustained by their environment, often symbolized by K). In EVOLVE, the maximum size is the one to which the population of adults may grow. If the survival and reproductive rates (and gene flow) are such that the population grows and exceeds the number you typed for maximum population, EVOLVE reduces the population to the size of the post-crash value. "Mortality" during the population "crash" is random with respect to genotype.

For example, if the number of young were 20,000, and the survival rates were all 50%, you would expect the number of adults to be about 10,000, twice the largest possible maximum population. If the post-crash value was 2,500, EVOLVE would multiply the number of each genotype by 2,500/10,000 to determine the actual number of adults.

You should choose these variables carefully to ensure that the population size is appropriate to the evolutionary situation you wish to model. For example, if you wish to study the effects of genetic drift in a small population of 10-20 adults, let the maximum = 20 and post-crash = 10; provided the population does not become extinct, it will remain within those limits. Setting maximum = 5000 and post-crash = 4000 would permit you to simulate a population with very little random change.

Appropriate choice of these variables is useful in other ways. Suppose you find that values for maximum and post-crash of 5000 and 4000 (coupled with high survival and/or reproductive rates) produce such rapid population growth that population crashes occur in every generation (the graph of total population size would be flat and fill the screen). If you wanted to get a better idea of how evolution affected rate of population growth, you could set post-crash to a lower value, such as 1000, and it would take longer for the population to exceed the maximum population of 5000.

Program Output

EVOLVE provides a number of ways of looking at experiments and their results. Each experiment has a Summary Window that is the control center for examining results. It contains a graph of genotype frequencies against time, a table of numerical results for each generation, and buttons that allow users to review the input parameters and generate graphs.

Information from any of EVOLVE's windows can be copied to the Mac's clipboard. From there the information can be Pasted into EVOLVE's Notepad, the Mac's scrapbook, a graphics program, a spreadsheet, or a word processor.

Additional graphs may be generated in two ways: By using the **Graphs** menu you may select standardized plots. You may construct graphs with any axes by clicking on the Make new graph button on the Summary Window.

The Copy Commands

Rather than produce reams of printed output, EVOLVE allows you to pick and choose what information you wish to look at and where and how you use that information. In this version of EVOLVE, Copy commands in the **Edit** menu are the only way information is moved. These commands can be used with any of the graph or data windows to move either graphic or numeric information.

Copy Window will save a bit-mapped copy of a window. The image includes every dot in the window, including its title and borders. Copy Window images have only 72 dots per inch (dpi) and do not have smooth lines when printed on a laser printer (unless the image is scaled down in size).

Copy Window Graph places a copy of the lines and text of a graph without the surrounding window. The image is not stored as dots, but as what is called a PICT, and can be printed with the full resolution of the printer you are using. When printed on a laser printer, Copy Window Graph images are free of the "jaggies" produced by Copy Window images (there may, however, be some problems with symbols and text in some situations).

Copy Window Data can be used to store numbers to the clipboard. The numbers are separated by [tab] characters and can be Pasted into most statistics and spread sheet programs where they can be subjected to more sophisticated analysis and graphing. They can also be copied into word processors if you need to report numerical results in a data table.

The Notepad

Each Summary Window has a Notepad that is a mini-word processor where you can type in your own notes, questions, hypotheses, predictions, and reminders. In addition, you can paste material that you have copied from EVOLVE's graphs and tables. The Notepad can be printed using the Print Notepad command in the **File** menu.

Using the Notepad is simple: Just bring a Summary Window to the front, click on the Notepad icon (▤) in the upper left-hand corner, and start typing and pasting. Click on the Notepad Window's close box to get it out of your way when you are doing other things.

The Variables You May Graph

It is often useful to look at the results of experiments from a variety of points of view, and EVOLVE is well equipped to let you examine its results in many ways. Where the actual list of variables is lengthy, they fall into a few distinct categories: time, numbers, and frequencies.

Here is a list of the variables that can be graphed:

> *Time:* from the start of the experiment measured in number of generations. This is so useful as the horizontal axis of graphs that it is the default.

The following group of variables are the actual number of adults in the total population or the number of a particular genotype or group of genotypes. By selecting combinations of the following, you can look either at genotypes or at phenotypes. The actual variables are:

> *Total Population:* the sum of the number of all genotypes, the total size of the population. This is important when you want to look at drift or population growth rate (mean fitness)

> *Size of •• Pop.*

> *Size of •◊ Pop.*

> *Size of ◊◊ Pop.*

Graphs of individual genotypes are useful at the introductory level to show what is happening to the population, and to show the Hardy-Weinberg relationship. Some

students ask for graphs of one or more genotypes when they have difficulty interpreting particular runs. Beyond these introductory uses, genotype graphs are probably of little value to most students. However, the steepness of the peaks, and the spacing between them, clearly illustrate that what ecologists call "intrinsic rate of natural increase" evolves over time and is a measure of the average absolute fitness of a population.

The last three population variables shown above can be used to graph the numbers of a phenotype. If, for example, • is recessive, then # •• would be the recessive phenotype and # •◊ + ◊◊ would be the dominant phenotype.

Each of these variables can be graphed on a logarithmic scale by selecting the `Log scale` box in the New Graph Window. A log plot enables you to look at numbers that vary greatly in size. For example, if the number of •• in an experiment went from 0 to 4000, a linear, non-log plot would not allow you to distinguish changes below several hundred individuals; a log plot would let you distinguish generations with only one or two individuals.

While actual numbers of individuals are instructive, they are often less useful than the relative frequencies of subgroups of the population. EVOLVE can also calculate the following as relative proportions of the population:

> *Frequency of* ••

> *Frequency of* •◊

> *Frequency of* ◊◊

Because evolution may be defined as "change in allele frequency," a graph of allele frequency is often the single best summary of what happened to the population. For EVOLVE, these allele frequencies are the most relevant measures:

> *Frequency of* •

> *Frequency of* ◊

The allele frequency graphs are always symmetrical about a horizontal line at 0.50 because the frequencies of the two alleles must always add up to 1.0 (if p symbolizes the frequency of the • allele, and q symbolizes the ◊ allele, then $p + q = 1$). The genotype frequency curves are more involved, but generally follow the Hardy-Weinberg relationship: $p^2 + 2pq + p^2 = 1$. Although this relationship does not apply precisely when selection, drift and/or gene flow operate, the actual frequencies are often fairly close.

Chapter 9. Theoretical Notes

The purpose of this chapter is to provide additional perspective on some of the theoretical issues behind the use of programs like EVOLVE. Many of the details, such as how to derive Hardy-Weinberg ratios, can be found in most textbooks and will not be covered here. Be sure to read the appropriate section of your text if this material seems unclear.

The Concept of an Equilibrium Population

The backbone of our theory of evolution in local populations, often called "microevolution," is the Hardy-Weinberg equilibrium concept. In its simplest formulation, the Hardy-Weinberg concept deals with a population with two alleles (• and ◊ in EVOLVE's notation) of one gene, that are present in frequencies p and q, respectively. Given such a population, four predictions may be made:

1. Frequencies of the two alleles will not change;

2. The ratios of the frequencies of the three genotypes (••, •◊, and ◊◊) will be p^2 : $2pq$: q^2 respectively, and will add up to 1.0;

3. The frequencies of the genotypes will not change;

4. If the genotype frequencies are not in the equilibrium ratios in one generation, they will reach equilibrium in the next.

Note that there are within-generation and between-generation predictions. Between generations, allele and genotype frequencies should not change (once equilibrium has been established). Within any one generation, the ratio of the genotypes should be p^2 : $2pq$: q^2 (•• : •◊ : ◊◊). Thus the Hardy-Weinberg "law" states that populations will remain in genetic equilibrium, that is, evolution will not occur!

Assumptions

You may find it odd that the foundation of our theory of evolution states that evolution will not happen, but the Hardy-Weinberg concept has a number of assumptions that invalidate it for most, if not all, natural populations. It strictly applies only when the following assumptions are true:

1. Individuals are diploid, reproduce sexually, and are identical with respect to such factors as age and sex;

2. Genes do not change from one allele to another;

3. The population is closed and alleles do not enter or leave;

4. Gametes are produced and combine randomly;

5. All zygotes survive and reproduce equally;

6. The population size is infinite or at least very large.

Evolutionary Forces

In essence, then, the conditions necessary for Hardy-Weinberg equilibrium tell us that for sexual species there are only six evolutionary forces (factors, processes, or phenomena that change allele or genotype frequencies):

1. Individuals are not equivalent (for example, the gene is sex-linked);

2. Genes change from one allele to another (*mutation*);

3. Genes move between geographic populations (*gene flow*);

4. There is *nonrandom mating* or gamete production and survival are not equal (for example, gametic mortality);

5. Individuals differ in their ability to survive and/or reproduce (*natural selection*);

6. Populations are of small size (*genetic drift*).

Obviously, populations that fit all of the above assumptions are rare, if they exist at all. However, some assumptions may not be important in particular populations, and it is often possible to relax others. For example, for all practical purposes populations on remote islands may be regarded as closed to immigration (assumption 3), and emigration may be considered as mortality (natural selection, assumption 5). Also, assumption 6 (infinite population size) can be disregarded if the population is fairly large (over a few thousand). Mutation (assumption 2) may be disregarded for most genes because mutation rates are too low to affect allele frequencies significantly in the short term. (Over the long haul, of course, mutation is critically important because it is the ultimate source of all genetic variation.)

Finally, although it is usually discussed for a gene with two alleles, the Hardy-Weinberg formula may even be extended to genes with more than two alleles by using the polynomial expansion. If there are three alleles in a population, their frequencies should add up to one ($p + q + r = 1$), and the frequencies of the genotypes would be ($p + q + r)^2 = 1^2$, or $p^2 + q^2 + r^2 + 2pq + 2pr + 2qr = 1$.

Importance of Hardy-Weinberg Equilibrium

Despite its limitations and qualifications, the Hardy-Weinberg concept has been useful in three major ways:

1. The Hardy-Weinberg concept makes *mathematical predictions* of what allele and genotype frequencies should be in the absence of evolution. It thus provides the "Null Hypothesis" needed for statistically rigorous tests of our ideas. See the BioQUEST statistics module if you are uncertain of this. If we have a way to measure those frequencies, and if they do not match the predictions, then we know that evolution is occurring. This obviously is important, because visible changes in populations happen so slowly.

2. The Hardy-Weinberg concept provides us with a *conceptual framework* for investigation. If evolution has occurred (that is, if allele or genotype frequencies deviate significantly from expected values) we know that one or more of the Hardy-Weinberg assumptions have been violated; we can proceed to determine which evolutionary force(s) have affected the population under study.

3. Finally, and perhaps most importantly, the Hardy-Weinberg concept provides the *foundation for a mathematically rigorous theory* of population genetics and for mathematical models of each of the evolutionary forces. Since most textbooks of population genetics illustrate these models in detail, we will not derive any here. These models may be used in turn as null hypotheses to evaluate the possibility that observed evolutionary changes are due to a specific evolutionary force.

Thus, although it paradoxically predicts no evolution, the Hardy-Weinberg equilibrium concept allows us to measure evolution, provides a conceptual framework for investigation, and serves as the foundation for the mathematics on which much of our theory of microevolution is based.

Models in Population Genetics

As indicated above, we can use the Hardy-Weinberg equilibrium concept to formulate mathematical models in the form of equations that hold all factors but one constant. We can then vary that one factor to investigate its effects on allele and genotype frequencies. Once we understand each of the forces by itself, we can combine it with other forces and examine interactions of more than one. In the absence of easy experimentation, population geneticists have turned to such equations for their experiments.

Much of our theory of microevolution is based on such models. Models permit us to explore the widest possible range of conceivable conditions — the theoretical maxima and minima of evolution. Thus the theory deals with "all possible universes"; experimental and observational studies must determine which universe actually exists.

Despite the importance of mathematical models in population genetics, they present problems. Elegant mathematical equations are intellectually stimulating and manipulating them can be instructive, but they require mathematical expertise and can be time-consuming. Also, mathematics are too often a barrier for many beginning students of evolution. The advent of calculators and computers has reduced computation time, but sometimes at the expense of a full understanding of the models, which are hidden from students. In addition, most mathematical models deal only with one or two of the evolutionary forces; it is difficult for students to get a feel for the way factors interact in populations. The abstractness of most such models makes it hard for students to really grasp what they mean — just what is a *selection coefficient* of 0.3? Finally, most of the simple mathematical models are deterministic. That is, they do not incorporate randomness or chance. Since chance plays a major role in evolution, deterministic models can be misleading.

EVOLVE is an attempt to avoid these difficulties. It is an intuitively simple yet realistic model because students have a better grasp of survival and reproductive rates than of selection coefficients. Depending on the values of input parameters, EVOLVE may incorporate any or all of three major evolutionary factors: natural selection, drift, and gene flow. Study of the fate of mutant alleles is also possible.

One of the things which people often have trouble understanding is the way in which chance events may affect evolution. Mutation, genetic drift, and stochastic (random) variation of environmental factors, such as weather and food supply, are examples of random factors in evolution. EVOLVE incorporates randomness in three ways.

1. The user specifies the average survival, reproductive, and immigration rates of the genotypes. Each time the program uses one of these parameters, a random-number generator is used to determine the actual value used. Thus, the environment of the model population may be regarded as varying slightly from generation to generation.

2. The size of the population may be made small enough for genetic drift to occur. Small sizes may be maintained over long periods of time, or may be temporary.

3. The pattern of random matings in a finite population, especially a small one, may result in deviations from the expected proportions of genotypes among offspring. You may find it helpful to make several runs with each set of data to get a feel for chance variations.

Calculating Fitness and Selection Coefficients

Despite the importance of measuring actual fitness and selection in nature, availability of the conceptual tools provided by the Hardy-Weinberg concept, evolutionary biologists have had difficulty actually collecting the data. Not until the advent of electrophoresis in the mid-1960's was there a method which gave clear data on the frequency of alleles in natural populations. However, there is some doubt about the evolutionary significance of electrophoretic variants, so even these relatively clear data

are difficult to interpret. Some scientists would rather look at morphological or life-history traits (such as length of teeth, age of first reproduction) that are of more obvious evolutionary value. Unfortunately, it is almost impossible to measure the genetic variability that underlies these phenotypes.

Aside from the difficulty of gathering the data, exactly how fitness and intensity of selection might be calculated from real-world data is a matter of considerable controversy. There is a large body of literature devoted to this discussion.

We might point out here that our difficulties with collecting and interpreting data in no way invalidate the scientific nature of our theory of evolution. What matters is that *in principal* the data could be collected. There are many examples in science of theories that could not be tested when they were proposed because of methodological difficulties. Wegner's theory of continental drift, for example, was proposed before 1920. We had to wait over 40 years until scientists had data on sea-floor spreading to indirectly support Wegner's hypothesis. We had to wait over 50 years before we could measure distances accurately enough to measure actual rates of continental movement.

What follows is a simplistic approach that will give you an idea of what is involved in measuring these important parameters. More importantly, this approach will help you understand the derivation of the concepts and what they mean. Here is an outline of the method:

1. Calculate observed frequency of each genotype in generation 1

2. Calculate expected frequency of each genotype in generation 2

3. Calculate absolute fitness of each genotype (R = obs/exp)

4. Calculate relative fitness of each genotype (W = fitness relative to the genotype with greatest absolute fitness)

5. Calculate selection coefficients of each genotype (s = 1 - W)

Here is an example that uses this approach. We will use the following data to calculate fitness and selection coefficients:

	Genotype			
	AA	Aa	aa	Total
Population in first generation (before selection)	4000	5000	3000	12000
Population in second generation (after selection)	3500	3000	1200	7700

1. Calculate observed frequency of each genotype in generation 2:

      Freq. of AA   = 3500 /7700 = 0.455

      Freq. of Aa   = 3000 /7700 = 0.390

      Freq. of aa   = 1200 /7700 = <u>0.156</u>

      Check: sum   = 1.001   (accurate to rounding error)

2. Calculate expected frequency of each genotype in generation 2. From the Hardy-Weinberg equilibrium, we expect the allele frequency in generation 2 to equal that in generation 1, and that the frequencies of the genotypes will be p^2, $2pq$ and q^2.

a. allele frequencies in generation 1 (= expected allele frequency in generation 2):

      Freq. of A = p = [2(4000) + 5000] / 2(12000) = 0.542

      Freq. of a = q = [2(3000) + 5000] / 2(12000) = <u>0.458</u>

      Check: sum   =   1.000

b. Expected genotype frequencies in generation 2:

      Freq. of AA = p^2 = $(0.542)^2$ = 0.294

      Freq. of Aa = $2pq$ = 2(0.542 * 0.458) = 0.496

      Freq. of aa = q^2 = $(0.458)^2$ = <u>0.210</u>

      Check: sum =   1.000

3. Calculate the absolute fitness of each genotype, $R_{genotype}$, selection) as the observed frequency divided by the expected.

      R_{AA} = 0.455 / 0.294 = 1.548

      R_{Aa} = 0.390 / 0.496 = 0.746

      R_{aa} = 0.156 / 0.210 = 0.743

4. Calculate the relative fitness of each genotype, $W_{genotype}$, as the absolute fitness divided by the highest absolute fitness.

$$W_{AA} \ = \ 1.548 \ / \ 1.548 \ = \ 1.000$$

$$W_{Aa} \ = \ 0.746 \ / \ 1.548 \ = \ 0.508$$

$$W_{aa} \ = \ 0.743 \ / \ 1.548 \ = \ 0.480$$

5. Calculate the selection coefficients of each genotype, $s_{(genotype)}$, as the relative fitnesses subtracted from one.

$$s_{AA} \ = \ 1.000 \ - \ 1.000 \ = \ 0.000$$

$$s_{Aa} \ = \ 1.000 \ - \ 0.508 \ = \ 0.492$$

$$s_{aa} \ = \ 1.000 \ - \ 0.480 \ = \ 0.520$$

There are a few points we should note about the concepts involved in these calculations. Intuitively, the selection coefficient, s, is the relative decrease in frequency due to natural selection. In this case, the Aa and aa genotypes are 49.2% and 52.0% less common than would be expected. By this definition, selection is not operating on the AA genotype.

Note that fitness and selection, as typically used by evolutionary biologists, are relative to the most fit genotype or allele in the population and are between 0.0 and 1.0. In every population at least one fitness coefficient must be 1.0 and the selection coefficient of that genotype must be 0.0. It is important to understand that these relative measures of the intensity of selection are intended to be used in discussing the relative changes in allele and genotype frequencies — which alleles and genotypes are increasing or decreasing in frequency.

Because these measures are relative, it is possible for a population to be declining in numbers at the same time that one or more alleles or genotypes have high relative fitness (indeed, one genotype *must* have a fitness of 1.0). Absolute fitness is the measure that tells us something about whether or not an allele or genotype can increase in numbers. If none of the absolute fitnesses is greater than 1.0, then the population will become extinct.

Strictly speaking, these calculations apply when data are collected in one generation and again at the same stage in the life cycle of the next generation (for instance, just after hatching, or at the start of mating). Data on the young of one generation cannot be used with adults of the next.

A similar approach may be used for calculating fitness and selection coefficients of alleles. Absolute fitness of alleles would be the ratio of the allele frequencies in generation 2 divided by their frequencies in generation 1. Relative fitnesses and selection coefficients would be calculated as shown above.

It is worth noting that science has a history, and that concepts developed at one time may not be relevant later. The concepts of relative fitness and selection coefficients were developed in the 1920s and 1930s because they made derivation of equations for population genetics easier. Given that it is difficult to measure actual coefficients in nature, this conceptualization of natural selection may in fact turn out to be unusable.

Limitations of EVOLVE

While it is important to understand the nature and value of the tools at your disposal, and simulation and mathematical models are powerful tools, it is also important to understand their limitations. In this section we try to clarify the nature of some of EVOLVE's limitations.

Among the most obvious of EVOLVE's simplifications is that we can look only at one gene with two alleles. Most evolutionarily important characteristics are influenced by many genes which often have many more than two alleles. Moreover, genes are not isolated — there are many types of interactions between genes. Even the interaction of two alleles at one locus may be affected by other loci; there is abundant evidence that advantageous alleles that are initially recessive can become dominant through the influence of such modifier genes. This over-simplification of the genetics of our model is perhaps the biggest conceptual problem with EVOLVE.

For example, one of the major questions in evolutionary biology is how the many diverse alleles seen in wild populations are maintained in natural populations. Using EVOLVE, you might find a half-dozen ways to maintain two alleles in a population over a long period of time. But, many evolutionary biologists don't feel these sorts of mechanisms would be sufficient and have extended their analysis to *three* alleles. This turns out to increase complexity significantly and to generate even more ways to maintain more than one allele in a population. The current version of EVOLVE cannot handle three alleles, although a future version will.

EVOLVE is also ecologically naive. In the real world, Carrying Capacity is rarely an absolute limit which is constant over time and which only influences a population when the limit is exceeded. Many limiting factors increase in severity as the Carrying Capacity is approached, and the Carrying Capacity itself changes over time. However, the J–shaped pattern of growth and crash illustrated in EVOLVE's population graphs does provide some measure of ecological reality as a consequence of the way limits are specified for genetic drift. In addition, the steepness of the Js and the spacing between them clearly illustrate that a population's intrinsic rate of natural increase evolves over time and is a measure of the average absolute fitness of a population.

EVOLVE'S concept of gene flow, the evolutionary consequence of the dispersal of individual organisms studied by ecologists, is an extremely simplistic one. For instance, the number of individuals immigrating is unlikely to be a constant: The source population(s) that produce immigrants will probably fluctuate; when the population is low it will produce fewer immigrants for the population modeled by EVOLVE. Also,

the size of EVOLVE's population should influence the number of immigrants that actually stay; if it is large, one would expect resources to be limiting and more potential immigrants to move on in search of more abundant food, shelter, or whatever. Finally, at high population densities when resources are in short supply, we would expect a higher proportion of the population to emigrate.

Despite its simplicity, EVOLVE does provide an entry into the world of theoretical population genetics. Although that world has evolved quite far from the initial simplicity of the Hardy-Weinberg equilibrium proposed eight decades ago, that view of nature still provides a useful starting point, especially for new students. To the extent that you develop some intuition and feel for the interaction of evolutionary forces, you may be misled into thinking you really understand evolution. Should you then have to apply that intuition in real evolutionary studies, you may find that it falls short. Nevertheless, we believe that you are better off working with EVOLVE, learning to think about allele frequencies and fitness, and designing and interpreting experiments, than merely reading a textbook.

Bibliography

Introductory Texts

Wilson, Edward O., and W. H. Bossert 1971. *A Primer of Population Biology.* Stamford, Connecticut: Sinauer Associates, Inc. Albeit old and dated, this is still the best, clearest, and shortest introduction to modeling in population biology.

Hartl, Daniel L. 1981. *A Primer of Population Genetics.* Sunderland, Massachusetts: Sinauer Associates, Inc. A more up-to-date, elaborate introduction to population genetics patterned after Wilson and Bossert.

Ayala, Francisco J. 1982. *Population and Evolutionary Genetics: A Primer.* Menlo Park, California: Benjamin Cummings. Another good introduction.

Spain, James D. 1982. *BASIC Microcomputer Models in Biology.* Reading, Massachusetts; Addison-Wesley Publishing Co. A superb, readable introduction to a wide variety of computer models in biology.

Full-Length Texts

Stansfield, William D. 1979. *The Science of Evolution.* New York: MacMillan Publishing Co. Although getting old as a survey of evolutionary biology, this text has a very clear derivation of a number of simple population genetic models.

Futuyma, Douglas J. 1979. *Evolutionary Biology.* Sunderland, Massachusetts: Sinauer Associates, Inc. An excellent theoretical and experimental survey of population genetics, ecology, and evolution.

Spiess, Eliot B. 1977. *Genes in Populations.* New York: John Wiley & Sons. A lengthy, more complete treatment of population genetics.

Wallace, Bruce 1981. *Basic Population Genetics.* New York: Columbia University Press.
    Another good survey of population genetics.

Advanced Texts

Crow, James F., and M. Kimura. 1970. *An Introduction to Population Genetics Theory.*
    New York: Harper & Row. This is a standard, mathematically-oriented text that
    is a classic in its field.

Wright, Sewell 1968-1978. *Evolution and the Genetics of Populations.* Vol. 1, *Genetic and
    Biometric Foundations;* Vol. 2, *The Theory of Gene Frequencies;* Vol. 3, *Experimental
    Results and Evolutionary Deductions;* Vol. 4, *Variability Within and Among Natural
    Populations.* Chicago, Univ. of Chicago Press. This series of volumes is the most
    advanced survey of the state of the art of population genetics.

Glossary

Allele   One of two or more differing forms that exist at one gene locus.

Allele frequency   A numerical measure of the commonness of an allele; the proportion of all alleles of a gene that are of a specified type. In a population of 20 ••, 30 •◊, and 50 ◊◊ individuals, the frequency of the • allele would be $((2 \times 20) + 30)/(2 \times (20 + 30 + 50)) = (40 + 30)/(2 \times 100) = 70/200 = 0.35$, or 35%.

Carrying Capacity   The number of individuals of a population that can be sustained by their environment over time; often symbolized by *K*.

Codominance   Pattern of inheritance where the heterozygote shows the phenotypic effects of both alleles. A good example is the allele for sickle-cell anemia. Homozygous Hb^AHb^A individuals are normal; homozygous Hb^sHb^s have severe, usually fatal sickle-cell anemia. Heterozygous Hb^AHb^s individuals have sickle-cell trait, a mild anemia.

Deme   A local population within which mating is random.

Diploid   The condition of having two sets of chromosomes (2N), one inherited from each parent. The alternative condition, *haploid* (N), occurs when there is only one copy of each chromosome. Gametes are haploid; zygotes are diploid.

Directional Selection   Pattern of selection that changes the frequency of an allele in a constant direction, toward either fixation or extinction of the allele.

Dominance   Pattern of inheritance in which an allele (the dominant) expresses its phenotypic effect even when it is in heterozygous condition with another allele (the recessive). If allele ◊ is dominant, then ◊◊ and •◊ individuals will have the same genotype, and •• homozygotes would have the recessive phenotype.

Electrophoresis   A technique for separating proteins in an electrical field. It enables small samples of proteins from wild organisms to be rapidly screened for differences in protein structure and hence in DNA.

Emigration   Movement of individual organisms out of a population.

Evolution   Broadly speaking, the origin of life and gradual change and diversification of living organisms over time. From the microscopic view of a single population, evolution is a change over time of the genetic and phenotypic composition of the population due to selection, drift, gene flow, nonrandom mating, and mutation. From the macroscopic view of many populations, evolution is the splitting of populations and their gradual divergence, coupled with extinction and other processes that give rise to the vast diversity of living forms that change over time.

Evolutionary Fitness   The contribution of an allele or genotype to the gene pool of subsequent generations.

Absolute F.   Ratio of the actual numbers of an allele or genotype in one generation divided by the numbers in a subsequent generation. Numbers greater than 1.0 indicate that the number of individuals with that allele or genotype is increasing.

Relative F.   Contribution of an allele or genotype to subsequent generations relative to alternate alleles or genotypes. Relative fitnesses greater than 1.0 indicate that the frequency of the allele or genotype is increasing.

Fixation Condition in which all members of a population are homozygous for one allele; alternative forms of the gene are extinct.

Gene   The fundamental unit of heredity; recognizable by the variant effects of different alleles on the phenotypes of the organisms carrying them; a segment of DNA at a particular location (locus) on a chromosome that affects some observable character(s) of the organism.

Gene frequency See Allele frequency.

Gene flow   Net movement of individuals from one population to another due to emigration and immigration.

Genetic drift   Random changes in allele frequencies due to sampling errors in finite populations; especially common in small populations where offspring are not a random sample of the parents' genes.

Gene pool   Abstract conceptualization of a population as the sum of all alleles of a given gene or of all genes together; subsequent generations are viewed as being drawn randomly from the gene pool.

Genotype   The specific combination of alleles present in an individual cell or organism.

Hardy-Weinberg equilibrium   Condition of stability in which allele frequencies of p and q will not change and the genotype frequencies will remain in a ratio of p^2:$2pq$:q^2. Hardy-Weinberg equilibrium requires that certain conditions be met (large population size, equal survival and reproduction of all genotypes, no differential gene flow, no differential mutation, random mating), and probably is rare in nature.

Heterozygote   An individual with two different alleles at one gene locus on the pair of chromosomes present in a diploid organism.

Heterosis The condition of "hybrid vigor:" the superiority of crossbred individuals over corresponding inbred individuals. Sometimes used to describe a condition where heterozygotes have higher fitness than homozygotes (overdominance for fitness). For example, in environments where malaria is prevalent,

heterozygotes for the sickle-cell anemia allele exhibit heterosis because they are more resistant to malaria than normal homozygotes and do not suffer the severe anemia of sickle-cell homozygotes.

Immigration   Movement of individual organisms out of a population.

Incomplete dominance   Pattern of inheritance in which a heterozygote shows a phenotype that is quantitatively intermediate between the phenotypes of the homozygotes. In four-o'clock plants, homozygous RR individuals have red petals, flowers of homozygous rr individuals are white, and heterozygous Rr individuals have pink petals.

Intrinsic Rate of Natural Increase   The difference between Birth Rate and Death Rate, often symbolized by r. Usually conceived of as the rate at which a population could grow in an environment with unlimited resources.

K   See Carrying Capacity.

Macroevolution The evolution of large phenotypic changes, usually large enough that the changed organisms may be regarded as a distinct, new genus or higher taxon.

Microevolution Evolution within a local population, usually resulting in small phenotypic changes; this is the level modeled by EVOLVE.

Mutation The process or event that produces an inheritable change in genetic material; at a single locus a mutation changes one allele to another. Also, the allele, chromosome, or individual that results from the process.

Mutation rate   The number of mutation events per gene per unit of time; for example, per cell generation.

Natural selection   The various processes that result in changes in the frequencies of genes or genotypes due to differences in the ability of their phenotypes to survive and reproduce.

Overdominance The expression of two alleles in a heterozygote of a phenotype that is outside the range of the corresponding homozygotes; see Heterosis.

Phenotype   The physical, physiological, biochemical, behavioral, or other properties of an organism that develop through the interaction of genes and their environment.

Pleiotropy   Condition in which a gene affects more than one phenotypic characteristic.

Population   A group of organisms of the same species that occupy a more or less well defined region.

r   See Intrinsic Rate of Natural Increase.

Recessiveness   Condition in which one allele expresses its phenotype only when it is homozygous; in heterozygous conditions, the recessive allele's effect is masked by the phenotype of the dominant allele. See Dominance.

Species   A group of organisms that actually (or potentially could) interbreed in nature and which is reproductively isolated from other such groups; the largest group within which gene flow could occur.