TABLE OF CONTENTS

.page

Chapter 1:
INTRODUCTION

EVOLVE is a computer program that allows you to experiment with evolution and to get quick results in a fashion that is impossible to do in any other way. You may control the starting population, overall population size, natural selection, pattern of inheritance and migration in a hypothetical population. By experimenting with EVOLVE you will develop:
1) a better understanding of evolutionary processes and their interactions,
2) a firmer grasp of some important concepts of Mendelian genetics,
3) some understanding of experimental design, and
4) of the use of models, plus
5) an appreciation for one of the many uses of computers in biology.

Perhaps we are belaboring the obvious to say that experiments in evolutionary biology are difficult -- you just cannot evolve something in a semester or even a lifetime! Nevertheless, this is an important point because it has had major effects on the study of evolution and has made the approach of evolutionary biologists 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. Experiments essentially are simple models of the real world that hold most factors constant, vary one or a few factors, and observe the results. In many areas of biology, experimental design has become a sophisticated and elaborate affair of choosing organisms, equipment, statistical methods, etc.

Evolutionary biologists can apply that approach only with difficulty. We can 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. However, observational tests of hypotheses are quite common in evolutionary biology and other historical sciences such as geology and astronomy.

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

HOW TO USE THIS MANUAL

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

We suggest that you first try the simple exercises 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. You could then skim through Chapter 3, 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 3 with more care. Chapter 4, MORE ADVANCED FEATURES OF EVOLVE, contains three more sample exercises to give you experience with all of the features of EVOLVE. You should do at least the first 2 exercises; the last may be omitted unless you will need to model a changing environment.

Part 2, EXPERIMENTING WITH EVOLUTION, contains a number of sample exercises that constitute experiments which both illustrate the capabilities of EVOLVE and demonstrate many features of evolutioary 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. 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 and, 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 and ADVANCED EXERCISES, are rather brief, for they are intended for advanced undergraduates and graduate students. Here the intent is to illustrate how to use EVOLVE as an experimental 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 more 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, Chapters 8 and 9, contain reference information on EVOLVE's menus and screen displays, setting up experiments and some more advanced topics. These chapters 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.
.page

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. Directions are italicized, notes and comments are in normal type; square brackets, [], denote special keys on the keyboard.

If the program is not already running, take the EVOLVE disk out of its protective cover, holding it by the label. Lift the latch of your computer's disk drive and insert the disk, oval cutout first and the label facing up. Close the door and turn on the computer. The light on the disk drive should come on as the program is read in from the disk. While the program is starting, continue reading.

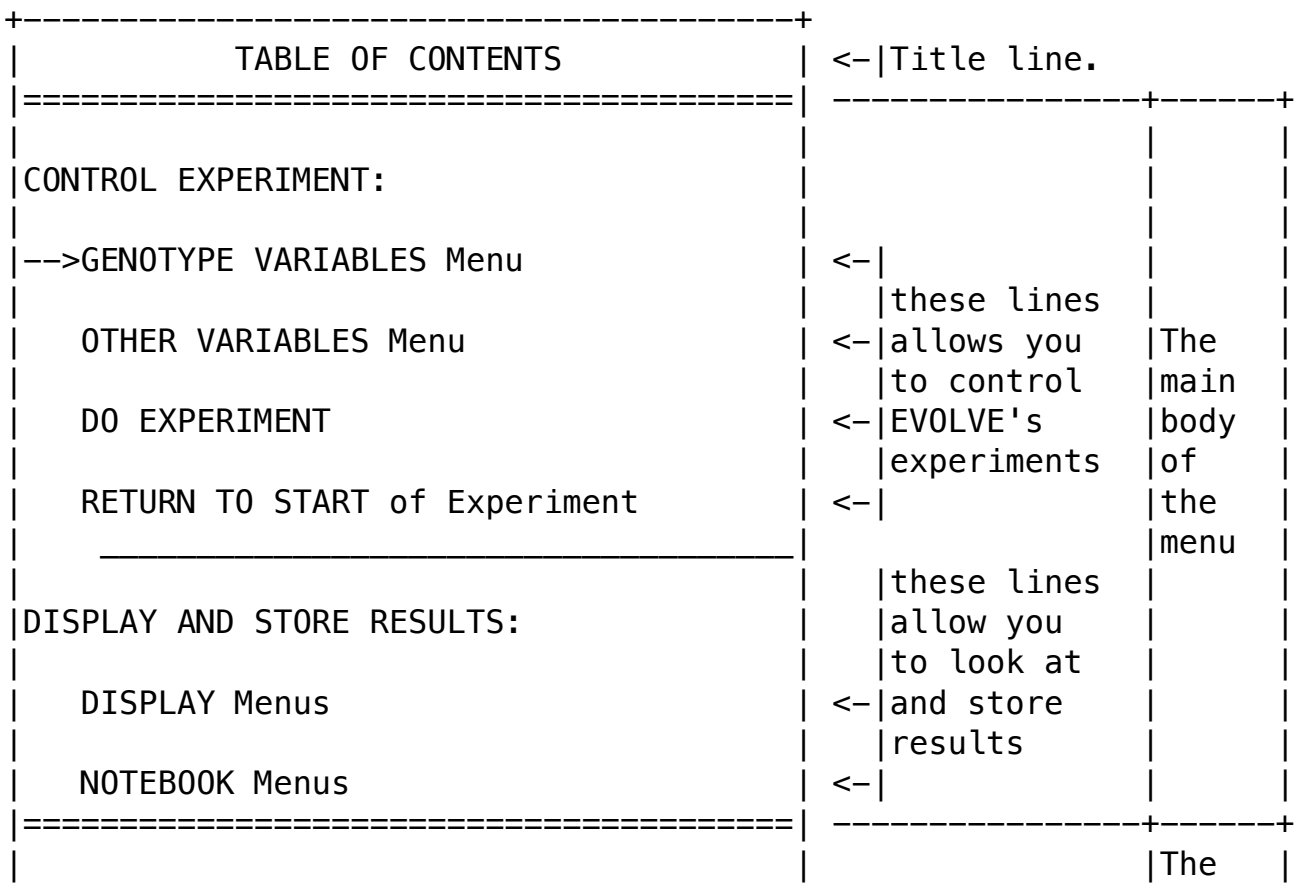
If someone else has been running the program, and you do not see the table of contents display (Figure 1-1), press the [ESC] key to return the program to the right place to start this tutorial.

EVOLVE is a "menu-based" program that always presents you with screen displays containing information about:
a) your options, what you may do at that point;
b) values of experimental variables (e.g., reproductive rates);
c) results of experiments.

Each menu or display consists of a menu ~title at the top, a ~main ~body containing the information, and a ~prompt ~area at the bottom listing actions you may take, along with messages and hints.

In this manual, special keys on the Apple keyboard are shown by words or symbols enclosed in square brackets, "[]". For example, [SPACE] indicates the spacebar, [RETURN] denotes the key labeled as such, [ARROWS] denote either of the two keys marked with left- and right-pointing arrows, [<--] and [-->]. On the screen, these special keys are shown as black images on a white background, what is called "inverse video".

By now, you should see a title page; press [RETURN] to view the first menu, shown in Fugure 1. This Table of Contents menu is your "home base." You can move to any other menu from here and you can return to this menu from any other menu or display by pressing the [ESC] key.



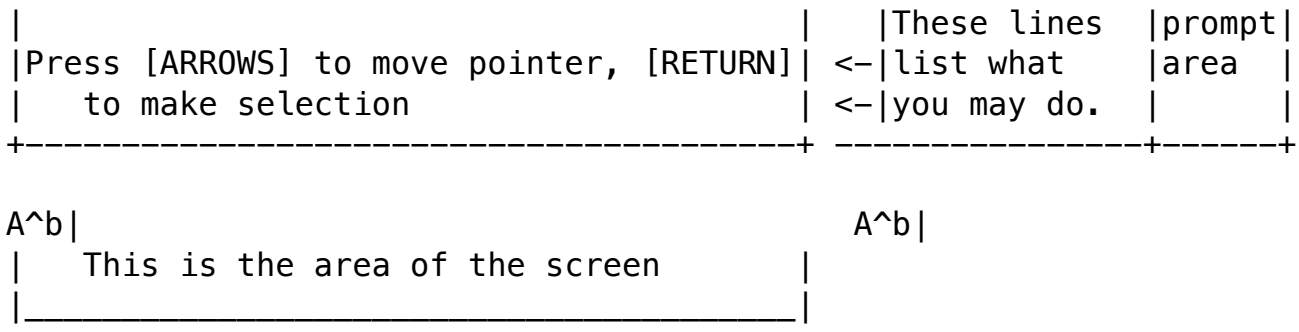


FIGURE 1-1. THE TABLE OF CONTENTS

The Table of Contents is typical of EVOLVE's menus. Note on the screen and in Figure 1-1, the title line, main body with choices of actions, and the prompt area. To the left of the screen is the pointer (symbolized by "-->" in the figure). You move the pointer using the [ARROW] keys on the computer keyboard, then select an item by pressing [RETURN]. Most of your communication with EVOLVE will be done with only three keys: [-->], [<--] and [RETURN].

Note that the main body of the table of contents is divided into two sections by a thin dashed line. The upper portion contains choices that will let you set up and run experiments. The lower portion has two choices that let you picture and store the results of experiments.

Press [<--] and [-->] repeatedly to see how the pointer behaves. Note that it does not move beyond the double lines; this area is the main body of the menu. You will now do a simple experiment to try out some of EVOLVE's menus.

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. To make the exercise concrete, think of EVOLVE as modelling a situation where a very large flock of migrating birds was blown onto an island and colonized it. We will look at a gene locus with two alleles symbolized by + and o. If individuals with the ++ genotype laid an average of 8 eggs, while the other genotypes (+o and oo) had an average of 5 eggs, the + allele would be an advantageous recessive. Assume further that the alleles had no effect on survival.

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 4000 organisms with a + allele frequency of 5%. Since we want the population to be in Hardy-Weinberg equilibrium, the initial population should consist of 10 ++ individuals, 380 +o heterozygotes, and 3610 oo homozygotes. (If you don't know about genetic equilibrium or how to calculate the numbers of each genotype, don't worry; these 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

Using the [ARROWS], move the pointer to the "RETURN TO START" line and press [RETURN]. The line will briefly be replaced with a message in inverse video, "Generation number reset to one." This will ensure that any previous results will be thrown away before you start your own experiment. Next, move the pointer to the "GENOTYPE VARIABLES" line and press [RETURN]. The screen will clear and the disk drive light will come on. After a short wait you will see EVOLVE's second menu, shown in Figure 1-2.

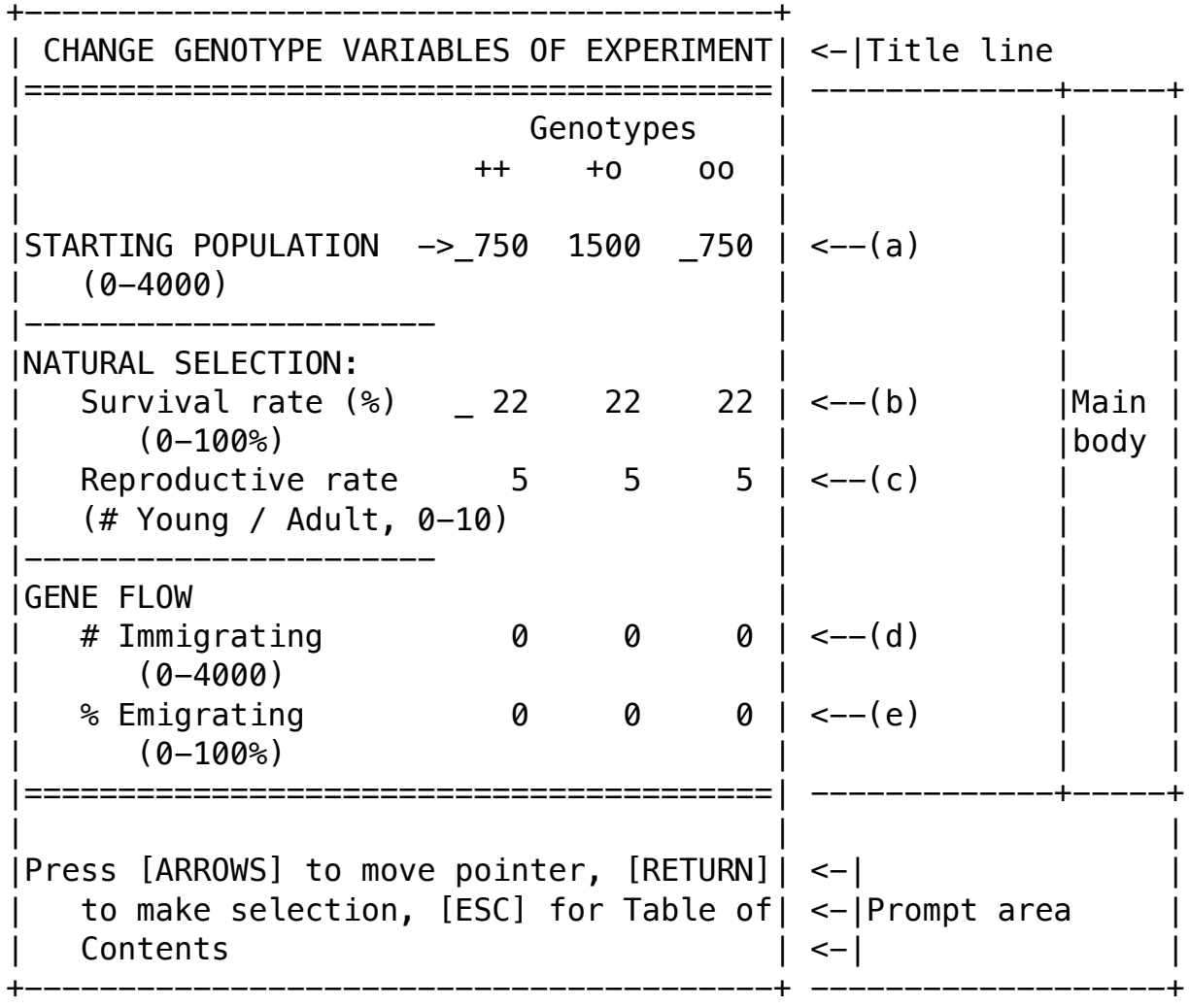


FIGURE 1-2. THE GENOTYPE CHARACTERISTICS MENU

EVOLVE models evolution of one gene having two alleles, + and o, and thus three genotypes, ++, +o and oo. This menu shows the relevant traits of each of the three genotypes. Proper choice of values for these genotype characteristics will allow you to establish a population and to determine the pattern of inheritance (dominance-recessiveness, etc.), natural selection, and gene flow that are the heart of EVOLVE's model of evolution.

The three columns headed ++, +o and oo contain numbers showing characteristics

of each genotype. We will briefly mention each of the lines, but defer lengthy discussion of the values to later; our objective now is to teach you to use EVOLVE, not to learn evolution.

The values displayed on the "STARTING POPULATION" line (a) show the number of adults of each genotype in the first generation. These values will probably not be the same on your screen and in the figure. The values of 750, 1500 and 750 in Figure 1-2 show a population containing 750 individuals of each homozygote, plus 1500 heterozygotes to make up a total population of 3000 adults.

Press [RETURN] and the pointer, with the number beside it, will disappear. Now type the number ten, followed by [RETURN]. Note that the digits appear on the left of the blank as you type them. They move to the right when you press [RETURN], and the pointer reappears beside the middle column. If you type the letters L or O, or anything except a number, nothing appears on the screen.

Whenever you need to enter a value into EVOLVE, rather than just select a menu item, this will be the pattern you follow: [RETURN] to clear the previous entry, type the new entry, and [RETURN] again to tell EVOLVE that you have finished. You must press the 2^und^d [RETURN] or EVOLVE won't remember your change. If you make a mistake before the second [RETURN], use [<--] to back up, then retype the entry.

Now press [RETURN] to remove the number of heterozygotes, enter 388 (yes, it should be 380, but type the extra 8 anyway) and then [RETURN] again to complete the entry. Now enter 3610 in the oo colum. This time you get a beep and a note in the prompt area at the bottom of the screen: "Total starting population cannot exceed 4000. Reduce last entry. Press [RETURN] to continue." Do so and the 3610 disappears. Now type a zero followed by [RETURN]. The pointer will now be on the first entry of line (b).

EVOLVE checks the values you enter and writes notes in the prompt area if it detects any mistakes. In this case, EVOLVE is complaining because it has room for only 4000 individuals in the starting population; we must go back and remove the extra 8 in the second entry. Press [<--] twice to move the pointer to 388, then [RETURN], 380, and [RETURN] again. Now enter 3610 into the oo column. Note that you may move back and forth in EVOLVE's menus and change entries as much as you wish.

The pointer should again be on the first entry on line (b). The three values on this line show the percentages of young of each genotype that survive to adulthood. Survival rates may range from 0 for a lethal genotype to 100 for a genotype that cannot die before reproductive age (note that the former is fairly common, while the latter is as common as hen's teeth). Now enter 22 into each space on line (b) to show that 22% of the young survive to adulthood. Since all genotypes have the same survival rate, we can regard the + and o alleles as not affecting inheritance of any characteristic related to survival.

Since we wish the + allele to be recessive and advantageous, enter 8, 5 and 5 into the ++, +o and oo columns respectively of line (c). The reproductive rates determine the average number of young that will be produced by adults of each genotype. These values may range from 0 offspring for a sterile genotype to 10 for a very fertile adult. Since the heterozygote has the same "phenotype" as the oo homozygote, the o allele is dominant over the + allele, and because the ++ homozygotes have a higher reproductive rate, the + allele is favored by natural selection.

Press [-->] to move to line (d). Lines (d) and (e) determine the amount and direction of gene flow into and out of the population. Make sure that all of these numbers are zero, so there will be no movement into or out of the population. We will not deal with these variables in our introductory exercises. The body of the Genotype Variables Menu should now look like Figure 1-3. If not, correct any differences. In a final look at this display, observe that the maximum and minimum values for each line are shown, so you don't need to remember them. When you are finished, press [ESC] to return to the Table of Contents. Again, note that the [ESC] key will always return you to the Table the Contents from any point in the program.

```
|=======================================|
|                          Genotypes    |
|                       ++    +o    oo  |
|                                       |
|STARTING POPULATION  ->_ 10   380  3610 |
|   (0-4000)                            |
|- - - - - - - - - - - -                |
|NATURAL SELECTION:                     |
|   Survival rate (%)   _ 22    22    22 |
|      (0-100%)                         |
|   Reproductive rate      8     5     5 |
|   (# Young / Adult, 0-10)             |
|- - - - - - - - - - - -                |
|GENE FLOW                              |
|   # Immigrating          0     0     0 |
|      (0-4000)                         |
|   % Emigrating          _0  ___0  ___0 |
|      (0-100%)                         |
|=======================================|
```
FIGURE 1-3. GENOTYPE CHARACTERISTICS FOR EXERCISE 1.

You should now be looking at the Table the Contents menu (Figur$ 1.1). Press [-->] to move the pointer down to the "Other Variables" menu, then press [RETURN]. After a short wait, you will see the display shown in Figure 1-4:

```
+---------------------------------------+
|  CHANGE OTHER VARIABLES OF EXPERIMENT  |
|=======================================|
```

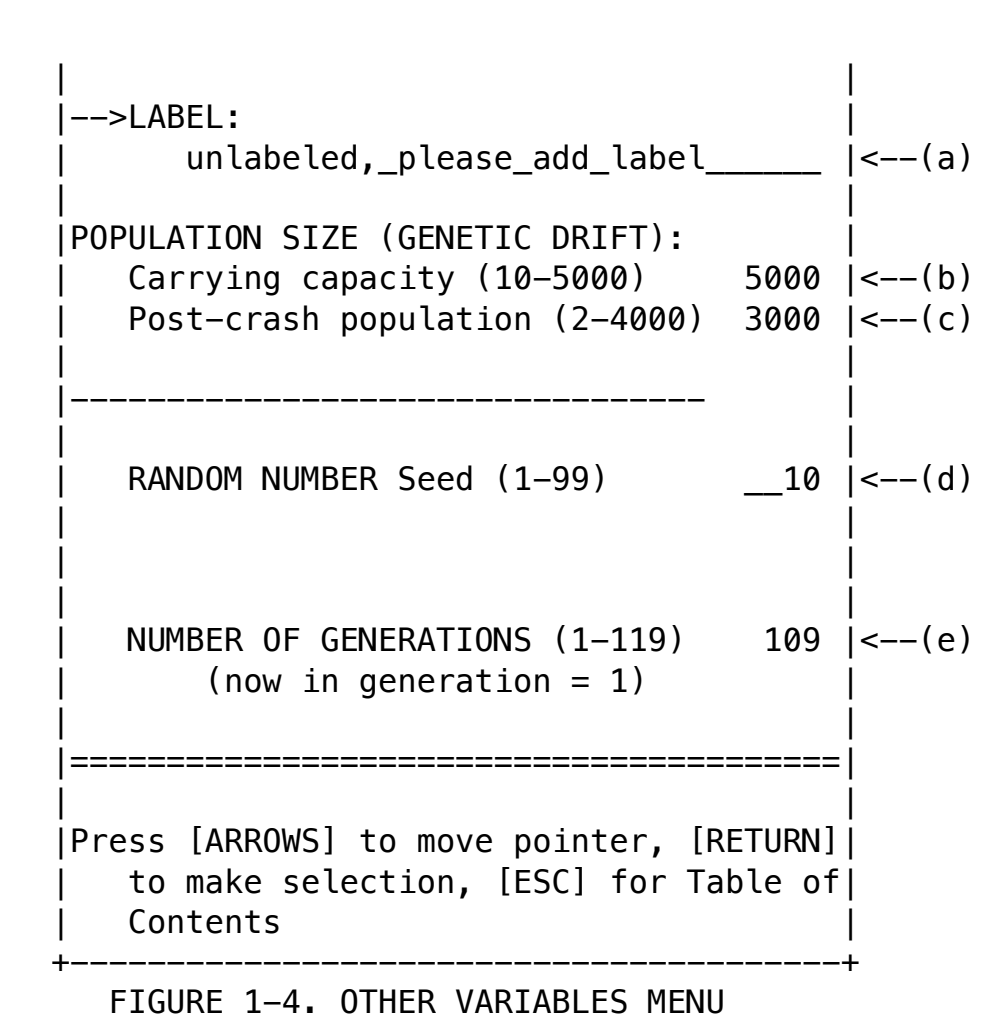


FIGURE 1-4. OTHER VARIABLES MENU

There are two types of items on this menu. Lines (b) and (c) determine the population size, and thus are an important part of any evolutionary experiment. The other items on this menu do not affect any of the evolutionary forces, but do determine how the computer manages the experiment -- such things as its duration.

Line (a), the label, is a title that you should enter to remind you of what the experiment is about. Move the pointer to the "LABEL" and press [RETURN], then type: EX. 1 - SEL FOR RECESSIVE [RETURN]. As before, if you make a mistake, press [<--] and correct the error. If you find a mistake after pressing [RETURN], you may go back and reenter the line. The label will appear as a title on the graphs and in the notebook if you choose to save the results. You will do a number of experiments and your graphs and notebook will get mixed up unless you use meaninful, unique titles to prevent confusion. Your labels should include relevant variables such as population size, type of inheritance, type of selection, random number seed, etc. The label for an experiment may include any combination of up to 33 numbers, letters or punctuation marks, including spaces.

Now press [-->] to move to line (b). Items (b) and (c) set upper and lower limits to the size of a growing population. Make sure these numbers are 5000 and 3000 respectively. The size of a population determines how much genetic drift will occur. If the survival and reproductive rates are such that the

population grows in size, then these limits become meaningful. If the survival and reproductive rates cause the population to decline, then it may decline to extinction; the "post-crash" limit is not a floor.

The "carrying capacity" on line (b) sets an absolute upper limit on the number of adults in the population. The "post-crash" value on line (c) determines the population size when the "carrying capacity" is exceeded. If the survival and reproductive rates are such that the adult population grows beyond the "carrying capacity" in line (b), then all genotypes are reduced so the total population is equal to the "post-crash" population size set in line (c). This process simulates, for example, the starvation that might occur if a population grew too big for its food supply.

Now move the pointer down to line (d). The random number seed plays a complicated role that is explained in detail later. At this point it is sufficient if you know that it affects the process of random mating and simulates variable environmental factors such as weather. Make this number 99 if it was something else.

Move the pointer to line (e). This number should be 109; if it isn't, you must go back to the Table of Contents menu, select the "RETURN TO START" option, then return to this menu. The number of generations tells EVOLVE how long to run the experiment. An experiment starts in generation 1 and may last as long as 109 generations, i.e., up to the 110^uth^d. If you should stop an experiment, the number on line (e) will tell you how many more you have to go, while the message below this line will show the number of the current generation. If you wish the experiment to continue for a specific number of generations, you may change this to any number less than that shown on line (e), and EVOLVE will stop automatically. We have now looked at all of the experimental variables; if the main body of your screen looks like Figure 1-5, you are ready to do your first experiment.

```
|=======================================|
|                                       |
|   LABEL:                              |
|      EX._1 - SEL FOR RECESSIVE    ____ |
|                                       |
|POPULATION SIZE (GENETIC DRIFT):       |
|   Carrying capacity (10-5000)     5000 |
|   Post-crash population (2-4000)  3000 |
|                                       |
|-------------------------------        |
|                                       |
|   RANDOM NUMBER Seed (1-99)       __99 |
|                                       |
|                                       |
|                                       |
|-->NUMBER OF GENERATIONS (1-109)    109 |
```

```
|       (now in generation = 1)          |
|                                        |
|========================================|
```

FIGURE 1-5. OTHER VARIABLES MENU FOR EXERCISE 1.

Doing the Experiment

Press [ESC] to return to the Table of Contents, then move the pointer down to the "DO EXPERIMENT" line (c on Figure 1-1); press [RETURN] to start the experiment. The screen will clear, the disk drive light will go on, and the display shown in Figure 1-6 will appear after a short time.

```
+----------------------------------------+
|     EX. 1 - SEL FOR RECESSIVE  _       |<-(a)
|1.0-|                           |       |
|0.9-|                           |  ++   |<-|
|0.8-|                           | ....  |<-|
|0.7-|                           |       |  |
|0.6-|                           |  +o   |<-|(b)
|0.5-|                           | ^u____^d |<-|
|0.4-|                           |       |  |
|0.3-|                           |  oo   |<-|
|0.2-|                           | ''''  |<-|
|0.1-|                           |       |
|0.0-|___________________________|       |
|      1  2  3  4  5  6  7  8  9  10     |
| Genotype Frequency vs Time (gen. x 10) |
|^u_____________________________________^d|
| Doing generation: xxx                  |<-|
|                                        |<-|(c)
|                                        |<-|
|                                        |
|========================================|
|                                        |
| Press [RETURN] to stop and restart     |
| experiment                             |
| Press [ESC] for Table of Contents      |
+----------------------------------------+
```

FIGURE 1-6. DISPLAY DURING AN EXPERIMENT

Let the experiment proceed for a short while and note that the number on the "Doing Generation:" line changes as the experiment runs. This display shows a graph of the frequencies of each genotype (on the vertical axis) over time (on the horizontal axis), along with other information about the progress of an experiment. As this first experiment proceeds, the lines will step across the graph and you should read the information below.

The title you entered on the Other Variables menu is shown on line (a) to

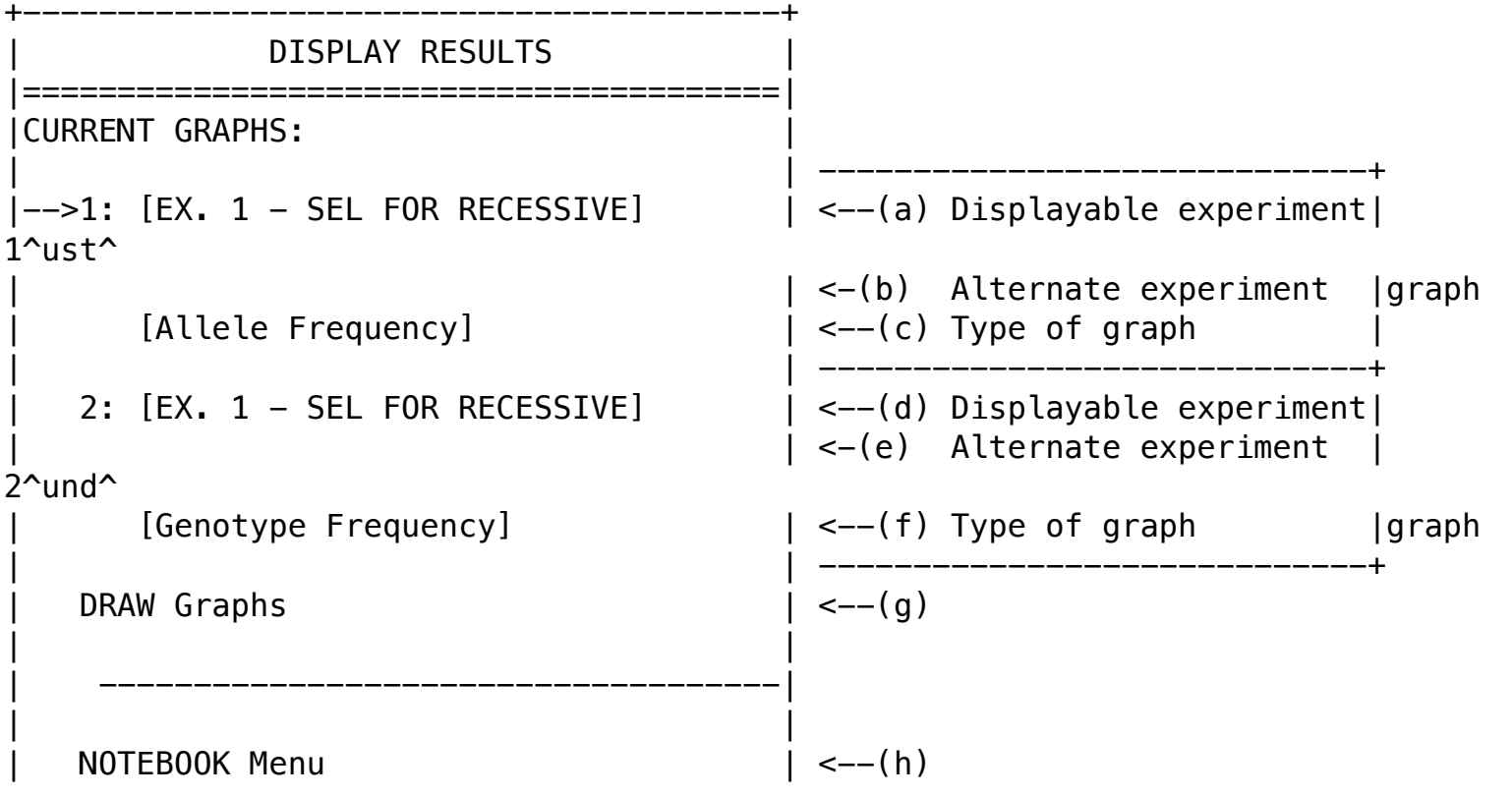
remind you of the experiment being run. On the right-hand side of the display, in the lines marked (b) in Figure 1-6, is a legend showing the line corresponding to each genotype. Below the graph is an area, labelled (c), that shows the generation number (xxx in Figure 1-6). If you stop the experiment, or if one of the alleles or the population becomes extinct, appropriate messages will be displayed here

If you wish to pause for thought or discussion, press [RETURN] once; press it again to resume. If you decide you do not wish to continue an experiment, press [ESC] to go to the Table of Contents. You may then display other graphs of the current results, and return to the start of the experiment to modify experimental variables.

During the 4 - 5 minutes it will take EVOLVE to run an experiment you should be trying to predict what will happen next on the screen. Predicting the results of experiments is an essential part of the scientific method and you should practice it continuously. When the experiment is finished, press [ESC] to go back to the Table of Contents.

Looking at Results

Use [-->] to move down to the "DISPLAY Menus" line, and press [RETURN]. After a short wait, you will see the display shown in Figure 1-7. If someone was using the disk before you, there may be another experiment's title on lines (b) and (d).



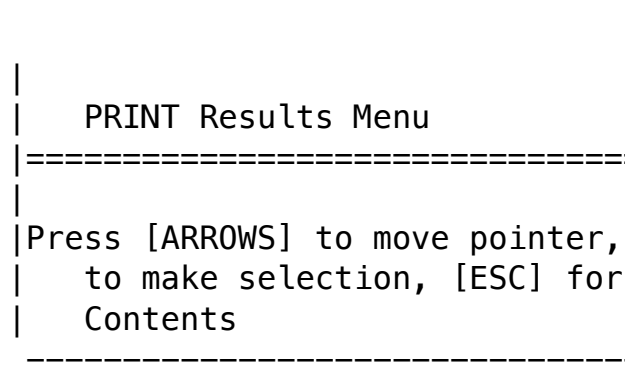
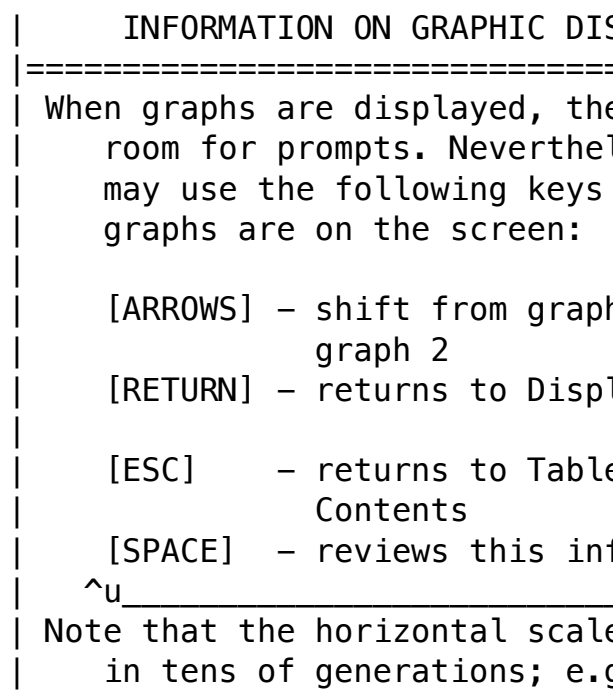


FIGURE 1-7. DISPLAY RESULTS MENU

This is a rather complicated menu because two graphs may be displayed at one time, and either or both graphs may be of any combination of 8 types of graphs of 2 different experiments. The labels and types of graphs that are ready to be shown are printed in inverse video on lines (a), (c), (d) and (f). The label of the alternate experiment (if any) is shown in normal video on lines (b) and (e). For now, don't worry about how to display a second experiment, that will be covered later.

If the title on line (d) is not the same a the one on line (a), move the pointer to line (d) and press [RETURN]. The titles on lines (d) and (e) will switch. If the types of graphs on lines (c) and (f) are not those shown in Figure 1-7, move the pointer to the line and press [RETURN]. The type of graph will change. Continue pressing [RETURN] until the correct type is shown.

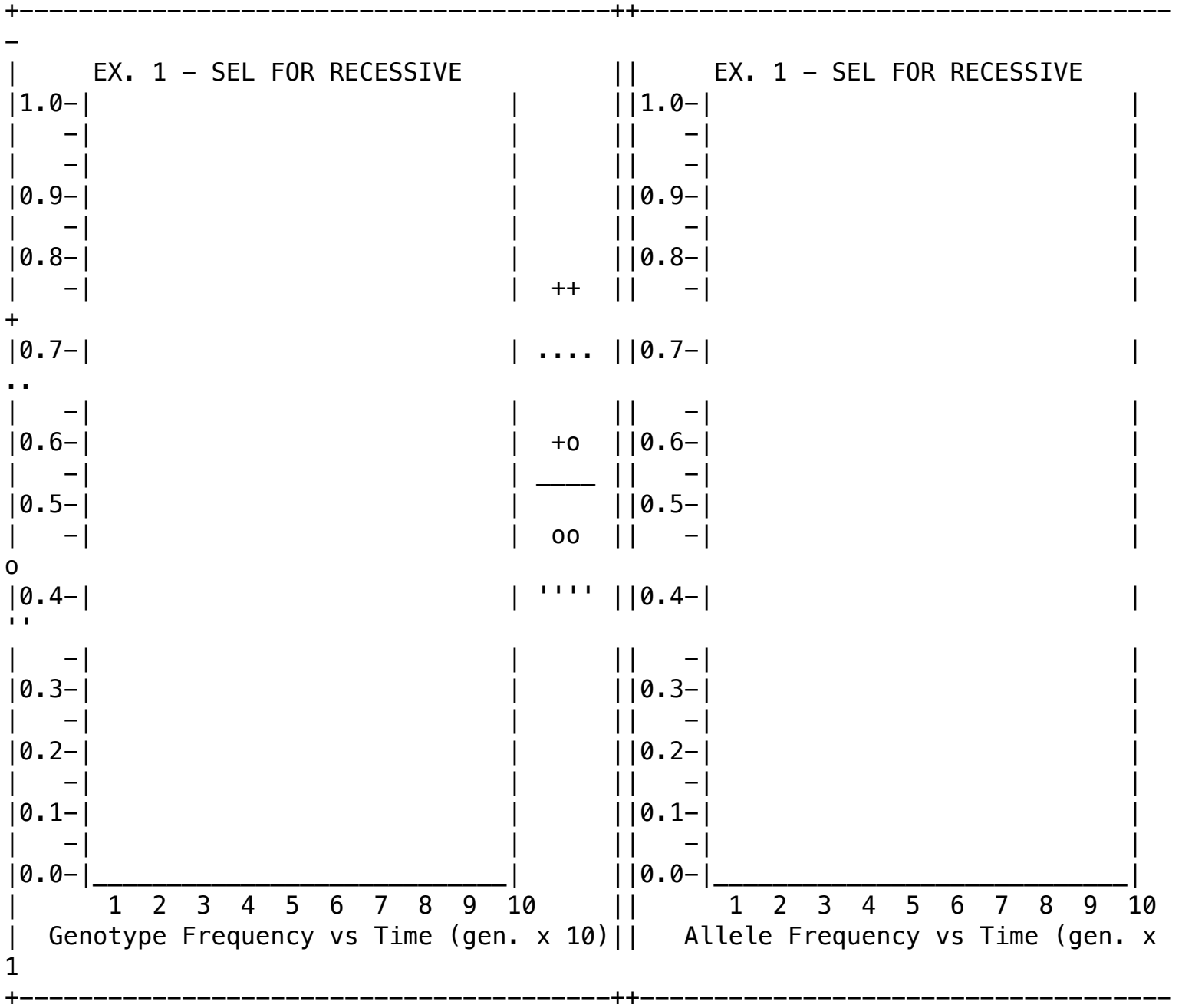
When the screen shows the same items on lines (a), (c), (d) and (f) as does Figure 1-7, move the pointer to "DRAW Graphs" (line g) and press [RETURN]. The screen will clear and you will see Figure 1-8 on the screen. Read these instructions, then press [SPACE_BAR].



```
|    ively.                              |
|========================================|
| Press [SPACE] now to view graphs.      |
|    Both graphs will be drawn, then the |
|    four commands listed above will be- |
|    come active.                        |
+----------------------------------------+
```

FIGURE 1-8. INFORMATION ON GRAPHIC DISPLAYS

After you press [SPACE], the screen will clear and Figure 1-9 will appear. After the graph is complete, press [-->] and you will see Figure 1-10. When the graph of allele frequency is finished, press [<--] and [-->] repeatedly to see their effect, then experiment with [SPACE].



1

FIGURE 1-9. GENOTYPE FREQUENCY GRAPH FIGURE 1-10. ALLELE FREQUENCY GRAPH

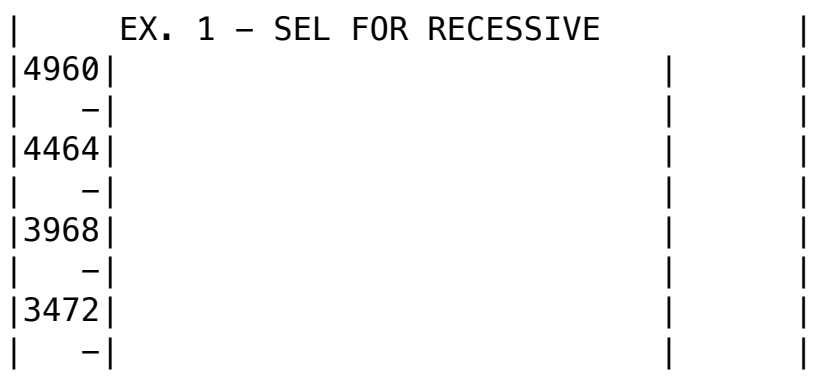
Note that these graphs are similar to the one displayed during the running of the experiment, although the vertical axis is larger to improve the resolution, and there is no room for prompts and other messages. There is a considerable amount of information to be gleaned from comparison of graphs such as these 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 some significant points.

Display the allele frequency graph (Figure 1-9). Note that the frequency of the recessive + allele was initially low (about 0.05, or 5% in generation 1) and climbed relatively slowly until about generation 40 (when its frequency was between 20 and 30%). It then rose more rapidly, becoming the most common allele after generation 50. After generation 70 its rate of increase slowed abruptly as its frequency approached 100%. On the other side of the coin, the dominant, but disadvantageous, o allele declined slowly from 95% in generation 1 to about 75% in generation 40. It then dropped precipitously until generation 70, when its rate of decline slowed. However, it become extinct after the 85^uth^d generation (it may look earlier because of the width and style of the line).

Now display the genotype frequency graph (Figure 1-10). This is complicated by the existence of three curves on the same graph, but you should note that the frequencies of the homozygotes cross at about 25% just after generation 50. The frequency of the heterozygotes rose slowly, peaking just above 50% at roughly the same time, then dropping precipitously to roughly the 70^uth^d generation. They started to level off, but became extinct in the 85^uth^d generation.

Press [-->] and [<--] repeatedly to switch rapidly between the two graphs. Note that this allows you to compare the two graphs and see that the heterozygotes peaked, and the homozygotes crossed, at about the time the allele frequencies crossed. There are theoretical, mathematical reasons for these relationships, which are worth discussing with fellow students or your instructor, but which we will not elaborate on here.

Press [RETURN] to go back to the DISPLAY RESULTS menu, change the type of graph 2 to "Total Population," then graph the results. The genotype graph will appear first, press [-->] to see Figure 1-11.



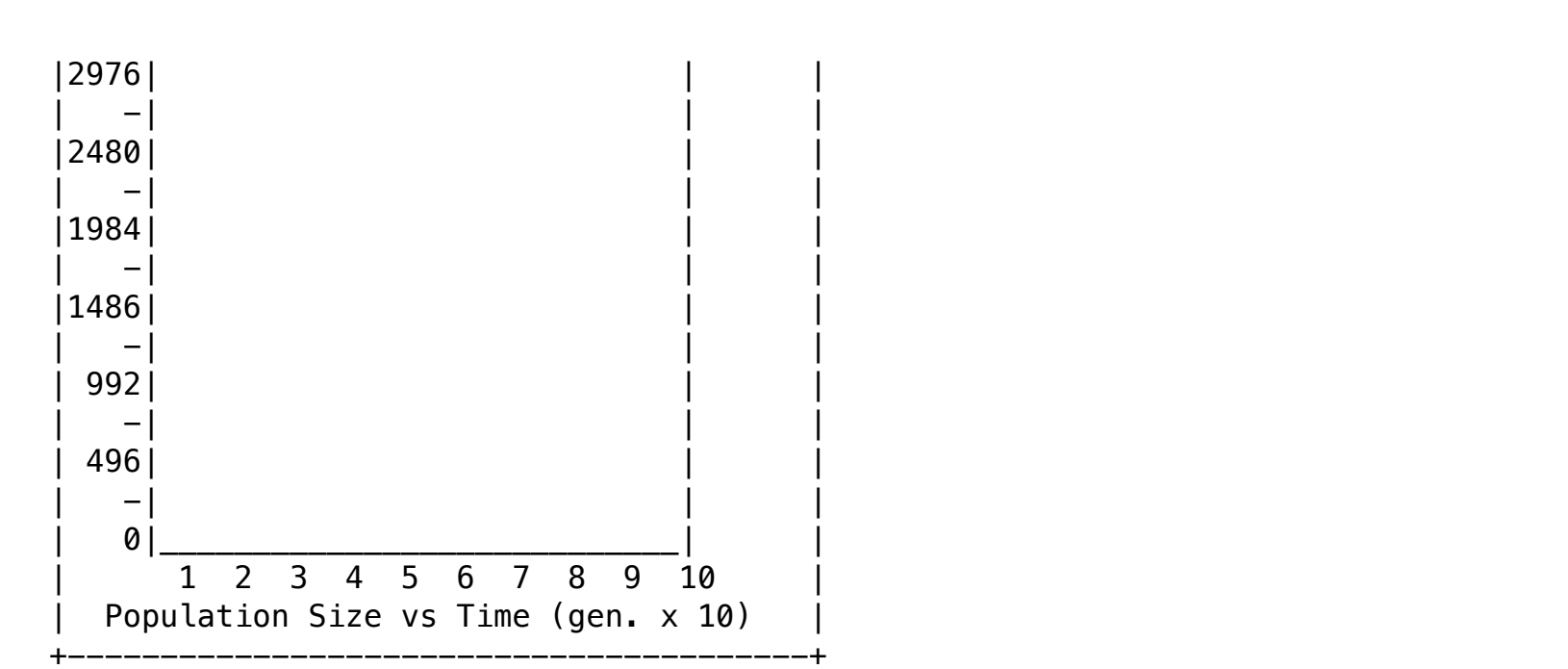
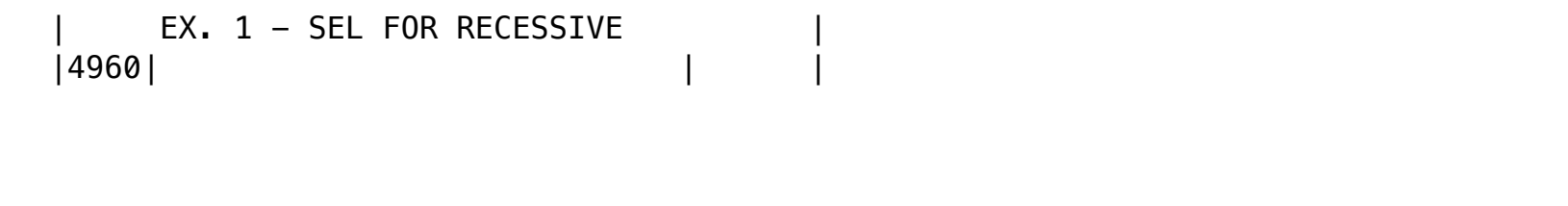


FIGURE 1-11. GRAPH OF TOTAL POPULATION SIZE

This graph differs from the first two in that it is a bar graph. Observe that the first bar goes up to the line labelled 3968. This is the closest point to 4000, the actual number in the initial generation, because the vertical axis is scaled from zero to the largest size attained during an experiment. The population size climbed to about 4900 in the 3^urd^d generation, then fell to the 2976 level in the 4^uth^d. Recall that the carrying capacity set in the "Other Variables" menu was 5000. During the experiment, the population of adults in generation 4 was greater than 5000, so EVOLVE "killed" enough to bring the numbers down to 3000, the "post-crash" size. The population continued this "saw-tooth" pattern of growth and crash, with the crashes becoming gradually more frequent until about generation 68. After that, the population size is a flat line at the "post-crash" size, for its growth rate was so great that it grew to over 5000 and crashed in each generation.

Now use the [ARROWS] to "flip" rapidly between the two graphs. Notice that as the frequency of the ++ genotype (the ones with the high reproductive rate) increased, the population crashes became more frequent. When the ++ individuals reached about 95%, the population crashed every generation.

[RETURN] to the Display Results menu, change the second graph to "Population of +o genotype" and draw the graphs. Again, switch rapidly between the genotype frequency and bar graph and note the relationship between frequency of heterozygotes and their numbers. Aside from the sawteeth caused by population crashes, the numbers of heterozygotes follow the curve of their frequency (see Figure 1-12). Notice that the bars stop when the o allele became extinct in generation 85.



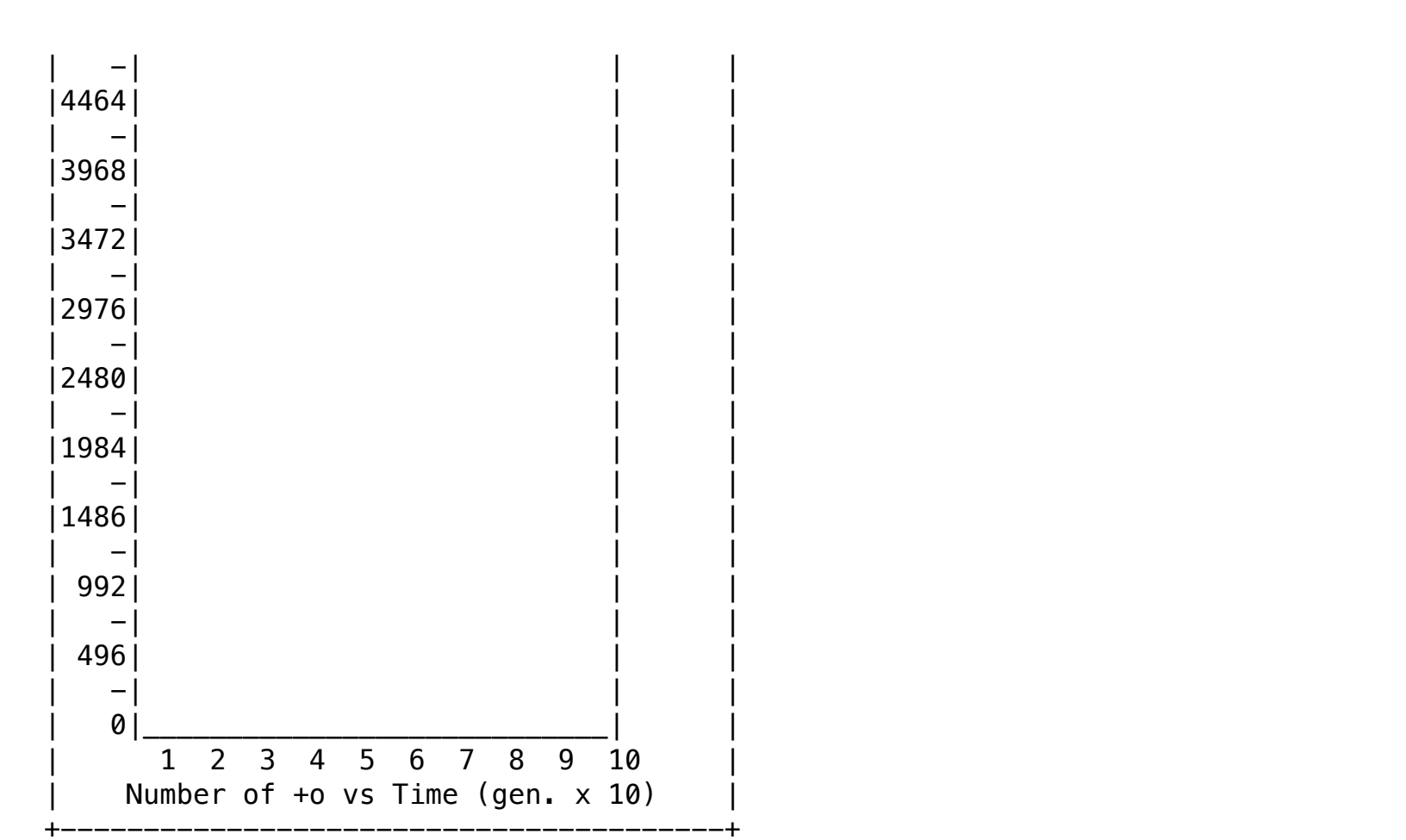


FIGURE 1-12. GRAPH OF NUMBER OF HETEROZYGOTES

In addition to the frequency and bar graphs that we have looked at so far, there is another type of graph -- the "selection coefficient" graphs for genotypes and alleles. These are difficult to explain and interpret, so we will defer them.

Leaving EVOLVE

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 two exercises in chapter 4. If not, leave EVOLVE as follows: Press [ESC] and return to the Table of Contents. Open the disk drive door and carefully remove the disk. Store it safely as described by your instructor, then turn off the computer.
.page

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 and 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 (the ecologists' "intrinsic rate of natural increase"). This chapter includes 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 frequency, and of 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.
.page

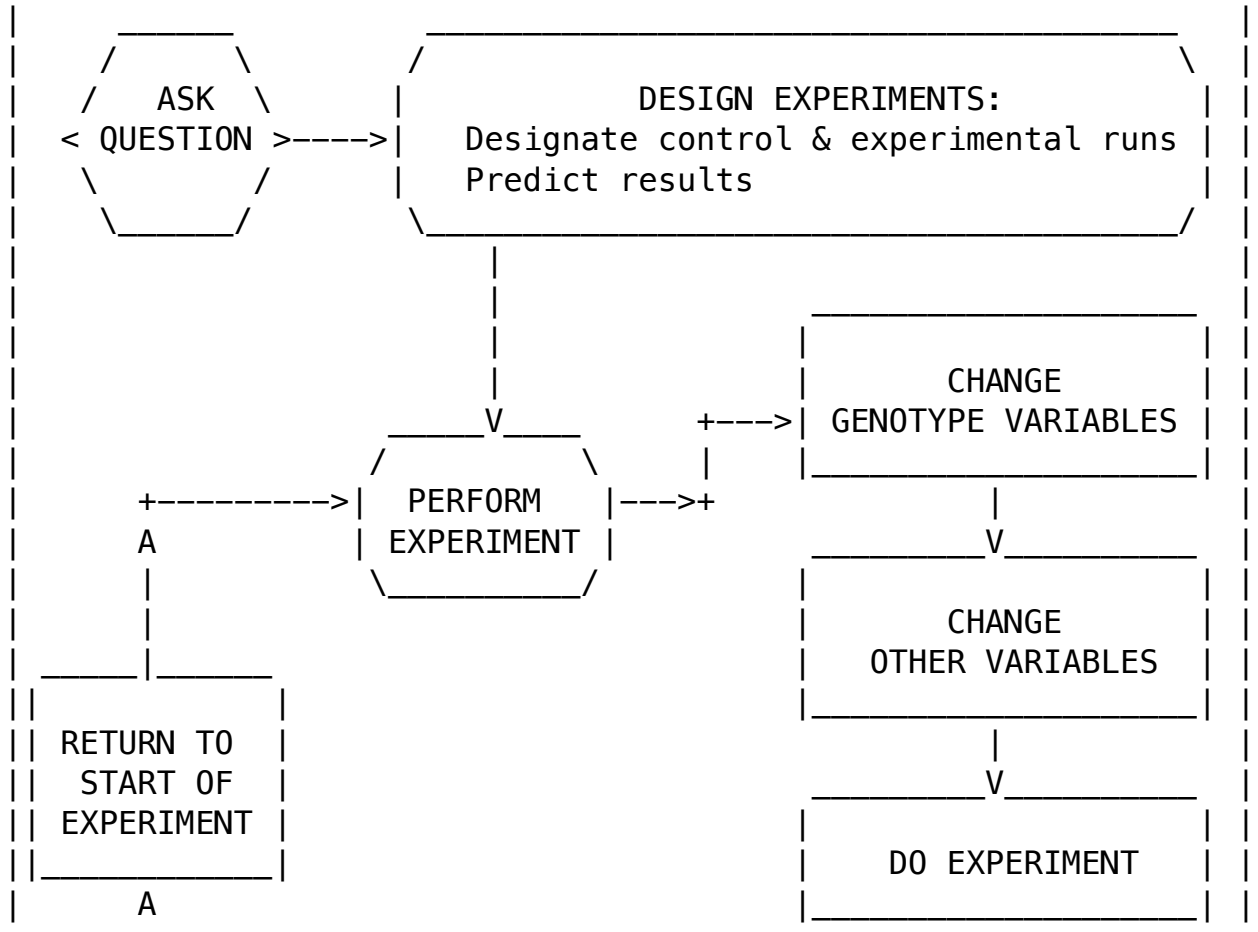
Chapter 4:
MORE ADVANCED FEATURES OF EVOLVE

In the last two chapters we introduced you to the most basic uses of EVOLVE -- setting up an evolutionary experiment, running it, and looking at graphs of 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 by 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 those experiments into proper context. Experiments should not be done haphazardly; they should be done in the context of a specific question (see Figure 4-1). The question should be rather specific and you should set up a minimum of two experiments to test it. One experiment should be designated the "control" against which the other(s) will be compared. You should also try to predict what the results will be in rather specific terms. Once the experiments are designed, you will set up the first one by using the Genotype Variables and Other Variables menus and then do the experiment. Once the results are in, you will probably want to store them in EVOLVE's Notebook (to be introduced in this chapter). You will then return to the Table of Contents, revise the evolutionary situation and do another experiment. Comparing the results of the experiments may lead you to revise your experiments or may give you enough information to answer your original question.

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



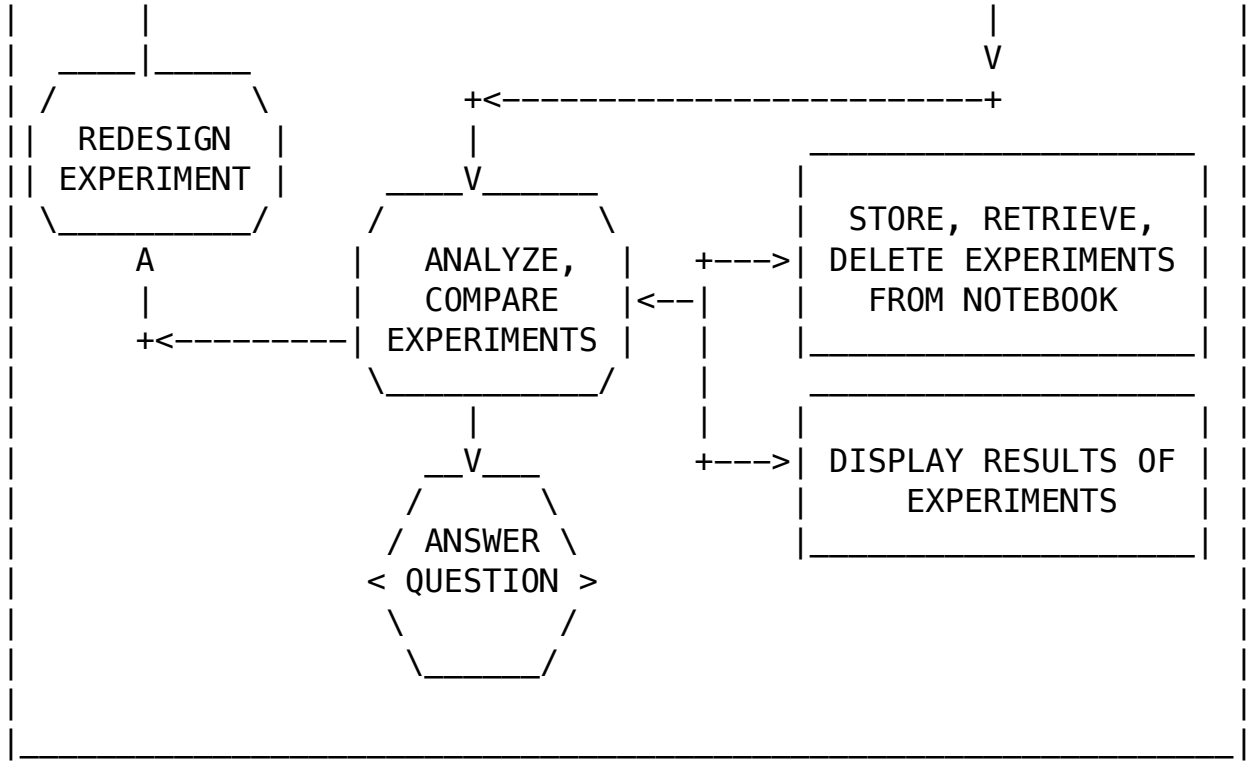


FIGURE 4-1. OUTLINE OF AN EVOLVE SESSION.
Rectangular boxes represent actions performed using EVOLVE menus, octagonal boxes show where user thought is involved.

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 a bit more specific. Since we have already done one experiment that provided data on natural selection in a population of 3000-5000, let us answer the more limited question, "Does selection for a recessive allele proceed differently in a small population of 30-50 individuals than in a large populatio of 3000-5000. 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 freqency graphs of the two experiments will be identical.

We will start with the experiment from chapter 2, store it in EVOLVE's notebook, modify it to keep the population small, and run the experiment. We will then compare the first and second experimental runs, and then make a third run to observe effects of changing the random number .

If you do not have EVOLVE running, start it up just as you did in Chapter 2. We will need to compare the results of Exercise 1 with those of Exercse 2. To compare results of two different experiments, you must first store them in EVOLVE's built-in notebook. Because many students may have used EVOLVE before you, always check to see that there is room to store your results. So, let us look in the notebook and see how to use it.

Using the Notebook

Starting at the Table of Contents, move the pointer to "NOTEBOOK Menus" and press [RETURN]. After a short pause while EVOLVE reads the disk, you will see the menu shown in Figure 4-1. This is perhaps the most confusing of EVOLVE's menus, so bear with our unavoidable lengthy discussion of it.

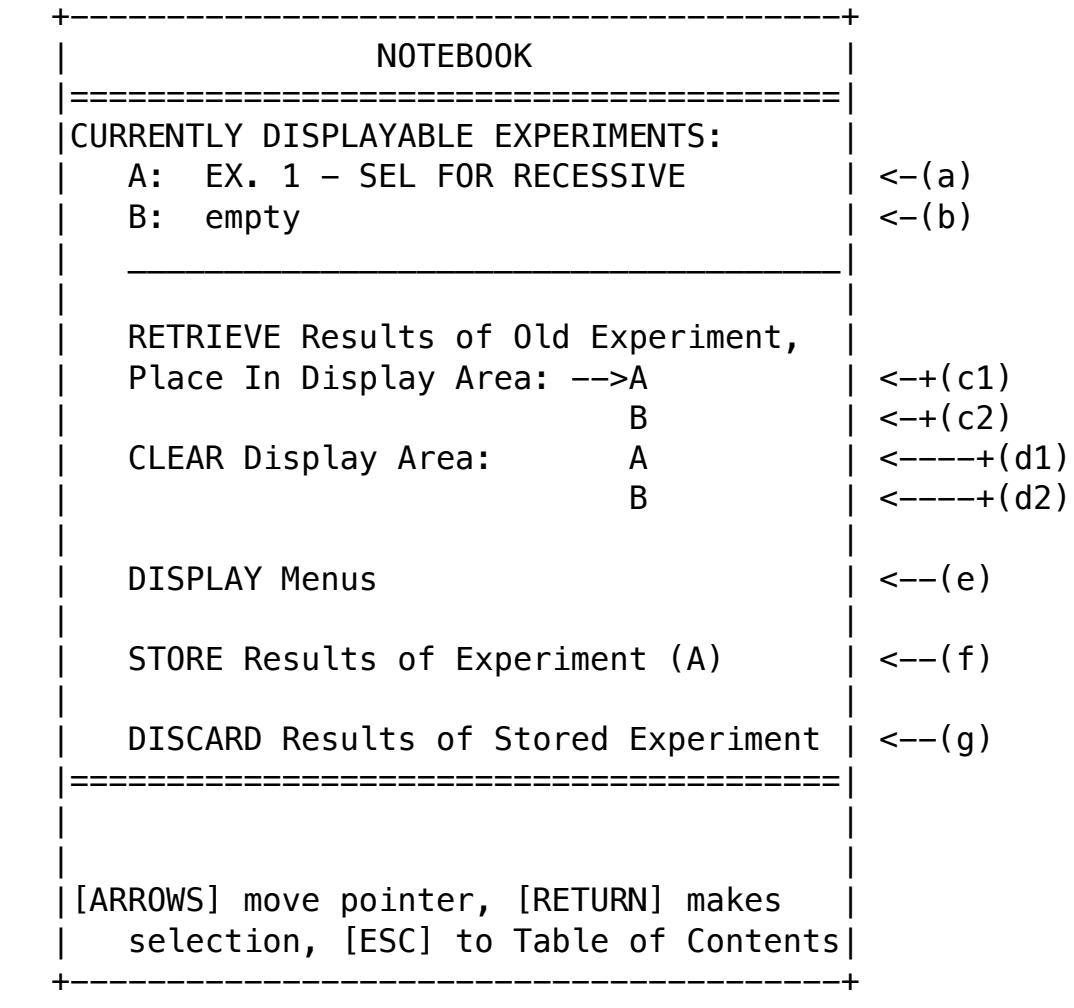


FIGURE 4-2. THE NOTEBOOK MENU.

EVOLVE can store up to 12 experiments in its notebook, but can display the results of only two at a time, just as you can look at only two pages of a book at one time. (EVOLVE has an advantage, however, in that you can tell it to move any two graphs next to each other for comparison.) The two displayable experiments are held in areas of the computer's memory that we will call "display areas" A and B. Lines (a) and (b) at the top of this menu show the

titles of the experiments that are held in these two display areas.

Listing notebook contents and discarding. -- Depending on how EVOLVE was used last, the titles on lines (a) and (b) of your screen may differ from those shown in Figure 4-2; if so, don't worry about it for now. Instead, let's look at a list of the notebook's contents and make sure there is space for us to work with. Select the "DISCARD Results" option by moving the pointer to that line (g in Figure 4-2) and pressing [RETURN]. You should now see something like Figure 4-3.

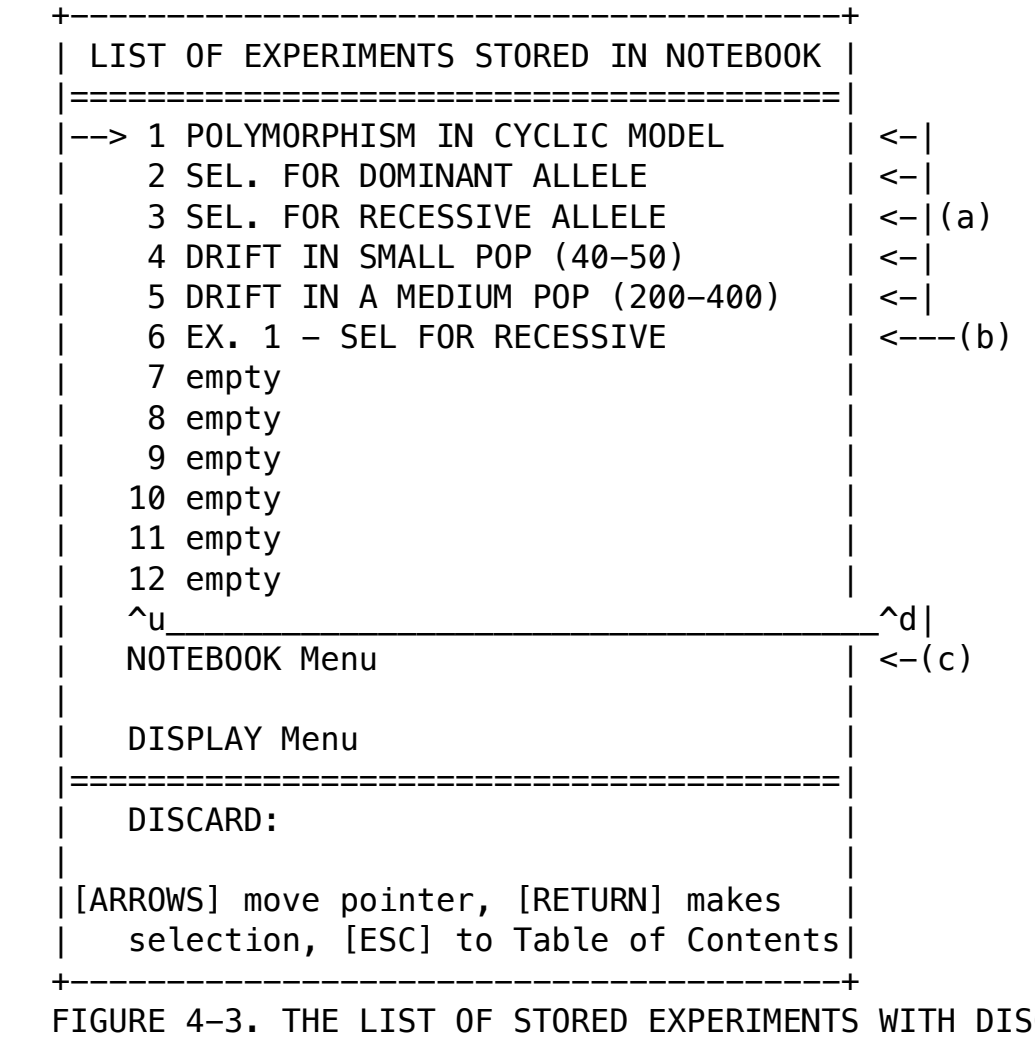


FIGURE 4-3. THE LIST OF STORED EXPERIMENTS WITH DISCARD OPTION.

The titles marked with (a) are just examples, you may not see anything like them. If you see the title of exercise 1 (as in line b, Figure 4-3), move the pointer to it and press [RETURN]. EVOLVE will display the number of the experiment next to "DISCARD" in the prompt area and ask "Are you Sure?" The pointer will be next to the word "yes." Double-check your selection. Press [RETURN] if it's correct; move the pointer and select "no" if you made a mistake. When you select "yes," EVOLVE will remove the title from the list and relabel that line "empty." If you should need to discard other results, you may continue to move the pointer and select as many titles as necessary. When you are finished discarding, move the pointer to "NOTEBOOK Menu" (line c) and press [RETURN].

Clearing display areas.-- We will be comparing results of two experiments in display areas A and B, so we must next make sure they are set up correctly. If display area B is not "empty," move the pointer to "CLEAR Display area: B" (line d2, Figure 4-2), and select it with [RETURN]. After a short pause, the second title will be replaced with the word "empty." Display area B is now clear for your use.

If display area A (line (a) in Figure 4-2), is "EX. 1 - SEL FOR RECESSIVE" then skip to the next paragraph. If not, select line (c1) to clear display area A. Pressing [ESC] will return you to the Table of Contents so you can repeat the experiment in Exercise 1, Chapter 2. Just reenter the data and do the experiment; you needn't look at the results again.

Storing and retrieving results. -- You should now have exercise 1 in display area A and display area B should be empty, if not, go back and correct them. Move the pointer to "STORE Results of Experiment (A)" and press [RETURN]. "[EXPERIMENT NOW BEING STORED]" will appear and the light on the disk drive will come on briefly. After the results are stored, move the pointer to line (c2), and press [RETURN] to retrieve the results just stored and place them in display area B. The list of stored experiments will reappear, with "RETRIEVE" in the prompt area instead of "DISCARD." Move the pointer and select the line with "EX. 1 - SEL FOR RECESSIVE" on it. EVOLVE will read the results from the disk and return to the "NOTEBOOK" menu.

Both display areas should now have the same title, "EX. 1 - SEL FOR RECESSIVE." You have seen how to use the notebook and organize the display areas, so you are ready to go back to the Table of Contents and do the next experiment. Press [ESC] to return to the Table of Contents.

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, i.e., the initial population, carrying capacity and post-crash size. We should also change the title to distinguish this from other experiments.

First, make sure that the generation number is reset to 1 by selecting the "RETURN TO START of Experiment" option. Now select the GENOTYPE VARIABLES Menu and check that it shows the same values as those in Figure 1-3. Next, change the starting population to 0, 4 and 36 for the ++, +o and oo genotypes respectively. This will result in an initial population that is midway between 30 and 50, which will be the post-crash and carrying capacity, just as 4000 was midway between 3000 and 5000 in Exercise 1. Also, the starting frequency of the + allele will be the same as in exercise 1.

All of this is important because we will be comparing the results of our second

run with those of the first run, i.e., 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 3/5^uth^ds of the carrying capacity) then we could not be sure that differences in results were due to differences in population size. The ratio of post-crash size to carrying capacity might have an effect.

Return to the Table of Contents with [ESC], then select the OTHER VARIABLES Menu. Change the label to "EX. 2 - SEL 4 REC (POP 30-50; 99)". This will reflect the fact that selection still favors the recessive allele, the population will be between 30 and 50, and the random number seed is 99. Now change the carrying capacity to 50, and the post-crash size to 30.

Press [ESC] to return to the Table of Contents, and do the experiment. As the lines march across the graph, try to predict what will happen next.

Once the experiment is complete, return to the Table of Contents and select the "DISPLAY Menus" option.

The display menu comes up with the latest experiment in "Current Graph 1," while the results of exercise 1 are ready for display in graph 2; both graphs are set to display genotype frequencies. Move the pointer to "DRAW Graphs" and press [RETURN] and [SPACE]. When the two graphs have been drawn, flip from one to the other, noticing the differences. While the graph of the small population is much more jagged than that of the large population, it generally follows the same pattern of relatively gradual change until roughly the 50^uth^d generation, then rapid change until the o allele became extinct in the 80^uth^d generation.

Use [RETURN] to get the DISPLAY RESULTS menu; change the graph type to allele frequency for both graphs and draw them. Again, the two graphs are quite similar in overall shape, although the small population's curves fluctuate a good deal more. As an aside, 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. In a similar fashion, compare the two experiments' other graphs (again, we will neglect to cover the allele & selection coefficient graphs because of their complexity).

In the graph of total population size from exercise 2, note the drop in population size to about 18 between generations 10 and 20. This was far below the post-crash size of 30 and was due to random fluctuations in survival and/or reproductive rates -- the computer equivalent of a period of drought, for example. Such changes have greater impact on a small population and show up more clearly when the vertical scale is 0-50 rather than 0-5000. Later you may

wish to print out summaries of exercise 1 to see if there was a similar decline in population size below the post-crash level.

In the graph of number of ++ homozygotes, notice that there are periods early in the experiment (e.g., generations 1-2, 11-18, 26-32) when this genotype did not occur in the population. Not until generation 46 did the ++ genotype return for keeps. Similarly, the oo homozygotes declined to zero, then reappeared before disappearing for good. Homozygotes, after all, may be generated by matings between heterozygotes when the rare allele is too rare for significant numbers of homozygotes to persist.

Printing Results on Paper

Our next experiment will compare Exercise 2 with another run that is identical except that it uses another random number. Before doing Exercise 3, however, we should note that EVOLVE can compare only two experiments at a time. So this might be a good time to show you how to preserve results on paper, how to obtain "hardcopy" in computer jargon. EVOLVE can print summaries of genotype numbers and frequencies, plus allele frequencies in each generation on virtually any printer. In addition, it can print copies of the various graphs on some printers; your instructor will have to tell you if this is possible on your equipment.

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, you may save your experiment in EVOLVE's notebook, take the disk to a computer that does have a printer, and print your results there.

If you do not wish to print a copy at the present time, skip the rest of this section and go on to Exercise 3.

=====> FINISH THIS SECTION - Print Ex. 1 <=====

EXERCISE 3. COMPARING RUNS WITH DIFFERENT RANDOM NUMBERS.

Our next exercise is designed to provide an illustration of how randomness may affect evolution. We will run an experiment which is exactly like Exercise 2, except for the random number seed.

Go to the Notebook, and store the results of Exercise 2. Clear display area B, then retrieve the results of Exercise 2 into it so both display areas hold experiment 2. Now return to the Table of Contents. Select the Return to Start of Experiment option, then go to the Other Variables Menu. Change the title to "EX. 3 - SEL 4 REC (POP 30-50; 90)" and then change the random number seed to 90. The net effect of these changes will be to repeat Exercise 2 with a different population and environment. This is a bit complicated, so we will discuss it briefly. But first, use [ESC] to return to the Table of Contents and

DO the experiment. While it is running, continue reading.

EVOLVE takes the "seed" you enter in the Other Variables Menu (90 in this case) and performs a series of mathematical operations on it to produce a "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 Genotype Variables Menu. This second number is also used as the seed for the third "random" number, and so on. Thus, 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, the wind direction and velocity, etc., 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."

The conclusion from all of this is that 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 (as we are doing now), you can assess the influence of chance. In essence, this amounts to running the same experiment again.

By now the experiment should be finished. Were the results what you expected?

If you need a reminder of the previous results, go to the Display Menu and compare the results with those of Exercise 2, or retrieve the results of Experiment 1 from the notebook into display area B and compare them with Experiment 3. Although the + allele was favored by selection, and the heterozygotes and homozygous ++ genotypes reached over 35% and 5% respectively, the advantageous allele became extinct! This phenomenon of random fluctuations of genotype and allele frequencies in small populations is an important, controversial evolutionary force called "genetic drift." Obviously, chance can have a profound influence on evolution.

EXERCISE 4. CHANGING THE EVOLUTIONARY SITUATION DURING A RUN

One of the simplifications often made in modelling evolution is to assume that the evolutionary forces are constant, i.e., that 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 as 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, are to be drawn randomly from a population having both alleles in equal abundance, and the alleles are assumed to be selectively neutral.

Go to the Table of Contents and return to the start of the experiment. Next, set up an experiment with the following data values (note that only 39 generations are to be done):

Title: Ex. 4 - The Bottleneck Effect

	Genotypes ++	+o	oo	Other variables	
Initial population:	1000	2000	1000	Carrying capacity:	5000
Survival rates:	22	22	22	Post-crash size:	3000
Reproductive rates:	5	5	5	Seed:	??
Immigrat., emigrat.:	0	0	0	No. generations:	39

Do the experiment. Note that the experiment stops in generation 39. Now return to the Table of Contents with [ESC], go to the Other Variables menu and revise them as follows:

```
 ___________________________
|                           |
|     Other variables       |
| ------------------------- |
| Carrying capacity:   50   |
| Post-crash size:     30   |
| No. generations:      2   |
|___________________________|
```

When the experiment stops in generation 42, go back and return the Other Variables to the following:

```
 ___________________________
|                           |
|     Other variables       |
| ------------------------- |
| Carrying capacity: 5000   |
| Post-crash size:   3000   |
| No. generations:     68   |
|___________________________|
```

Go to the Display results menu and 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 3000 and 5000 strong; after the period after the 2-generation crash, however, the allele freqencies "drifted" away from 0.50. As the population grew back, however, 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 each generation was a sample from a smaller previous population.

====> add picture of output? <=====

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 chapter lists 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 providemor
.page

Chapter 5:
ELEMENTARY EXERCISES

This chapter is 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 design experiments. The chapter is

designed to provide 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 is the heart of the chapter and provides a series of exercises that guide you through experiments with single evolutionary processes; the final section illustrates how pairs of evolutionary forces interact.

FUNDAMENTALS

Questions 1-6 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, for the latter will assume that you know the concepts involve with these first exercises. Your instructor may suggest that you treat this as a take-home quize 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 o allele frequency of 0.63; write the number of each genotype in the spaces below:

No. ++ = _____ No. +o = _____ No. ++ = _____

2. Consider an initial population of 448 ++ individuals, 1238 +o individuals, and 855 ++ individuals.

a. What are the frequencies of the alleles?

Frequency of + = _____ Frequency of o = _____

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. +o = _____ 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 (hatching, germination, etc.) to reproductive age. Reproductive rates are measured as the

average number of young born per individual of each genotype.

a.

"Phenotype"	Genotype ++	Genotype +o	Genotype ++
Surviv. rate	25%	25%	20%
Reprod. rate	3	3	3

b.

Genotype ++	Genotype +o	Genotype ++
50%	30%	30%
2	3	3

c.

'Phenotype"	Genotype ++	Genotype +o	Genotype ++
Surviv. rate	40%	40%	58%
Reprod. rate	1	2	2

d.

Genotype ++	Genotype +o	Genotype ++
12%	18%	24%
9	9	9

e.

'Phenotype"	Genotype ++	Genotype +o	Genotype ++
Surviv. rate	60%	75%	65%
Reprod. rate	2	1	1

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 o 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 first two exercises are set up for you to illustrate what is required and

to give you an introduction to the whole process. In particular, note the way the parameter values are set up, 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 EVOLE. 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, but change the random number seed to see what differences this makes. 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 and in which no evolutionary forces are at work?

This exercise is laid out in rather complete detail to help you develop an understanding of how to set up your own experiments, read and interpret the results, and draw conclusions. Later exercises will build on your developing expertise and will be spelled out in less detail

Experiment and Prediction: A population in obvious disequilibrium would be one consisting entirely of heterozygotes. 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^u2^d for ++, 2pq for +o, and p^u2^d for oo.

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
Seed:	(enter a 2-digits of your choice)
Initial Population:	0000 ++, 4000 +o, 0 oo
Carrying Capacity:	5000
Post-crash Pop. Size:	1000
Survival Rates:	26 ++, 26 +o, 26 oo
Reproductive Rates:	4 ++, 4 +o, 4 oo
Number of Generations:	10

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

a. What is the frequency of the o 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? ++: _____ +o: _____ oo: _____

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

d. Why is the carrying capacity set at 5000 and the post-crash size to 1000?

Results:
Fill in the following data table from the summaries printed by EVOLVE (if you do not have access to a printer, record the data as best you can from the graphs 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.

Obeserved Frequencies

	Genotype frequency			Allele frequency	
Generation	++	+o	oo	+	o
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

	Genotype frequency			Allele frequency	
Generation	++	+o	oo	+	o
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.

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 o allele equals 25%, the initial genotype frequencies are in Hardy-Weinberg equilibrium and the population totals 4,000 individuals. (Let 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 grows slowly and in which there is no selection. But 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.

Initial Population	:	_____++, _____+o, _____oo, _____Total
Carrying Capacity	:	_____
Post-crash Pop. Size	:	_____
Survival Rates	:	_____++, _____+o, _____oo
Reproductive Rates	:	_____++, _____+o, _____oo

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 and evolutionary "progress." 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 &/or reproductive rates of the two phenotypes should show that evolution (i.e., 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 (o)
Seed	:	(choose a 2-digit number)
Initial Population	:	1711 ++, 278 +o, 11 oo
Carrying Capacity	:	5000
Post-crash Pop. Size	:	3000
Survival Rates	:	20 ++, 24 +o, 24 oo
Reproductive Rates	:	6 ++, 6 +o, 6 oo
Number of Generations	:	109

2nd run:

Title	:	7B. STRONG SEL FOR DOM ALLELE (o)
Seed	:	(use the same number)
Initial Population	:	1711 ++, 278 +o, 11 oo
Carrying Capacity	:	5000
Post-crash Pop. Size	:	3000
Survival Rates	:	20 ++, 34 +o, 34 oo
Reproductive Rates	:	6 ++, 6 +o, 6 oo
Number of Generations	:	109

Why should the seed be the same in both runs?

Results:
Fill in the following tables from the graphs produced by EVOLVE. Note that, because of rounding, the graphs are only accurate to +^b_2%, but that is close enough for our purposes. 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.

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

Run with strong selection

	Change in			
	Genotype frequency			Frequency of
Generation	++	+o	oo	+ allele

1 -> 5	_____	_____	_____	_____
6 -> 10	_____	_____	_____	_____
11 -> 15	_____	_____	_____	_____
16 -> 20	_____	_____	_____	_____
21 -> 25	_____	_____	_____	_____
31 -> 35	_____	_____	_____	_____
51 -> 55	_____	_____	_____	_____
71 -> 75	_____	_____	_____	_____

Run with weak selection

	Change in			
	Genotype frequency			Frequency of
Generation	++	+o	oo	+ allele
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 data from the graphs of population size versus time:

	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:

Conclusion:
Finally, summarize the results, i.e., compare the evolution of dominant, advantageous alleles under strong and weak selection pressure.

EXERCISE 8. Does the evolution (i.e., 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, have your instructor initial it):

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 with a different random number seed -- you would 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 (o) is recessive:

```
   Title                : 8A. SEL FOR RECESSIVE ALLELE (o)
   Seed                 :   (choose your own 2-digit number)
   Initial Population   :   1711 ++,  278 +o,   11 oo
   Carrying Capacity    :   5000
   Post-crash Pop. Size :   3000
   Survival Rates       :     20 ++,   20 +o,   24 oo
   Reproductive Rates   :      6 ++,    6 +o,    6 oo
   Number of Generations:    109
```

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, have your instructor initial it):

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
Seed	:	(choose your own 2-digit number)
Initial Population	:	1711 ++, 278 +o, 11 oo

```
   Carrying Capacity    :   5000
   Post-crash Pop. Size :   3000
   Survival Rates       :     20 ++,   24 +o,   20 oo
   Reproductive Rates   :      6 ++,    6 +o,    6 oo
   Number of Generations:    109
```

Prediction
(enter your own, have your instructor initial it):

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 ++,   0278 +o,   1711 oo
```

Prediction
(enter your own, have your instructor initial it):

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

```
   Title                : 10C. SEL FOR HETEROZYGOTES
   Initial Population   :   11 ++,  278 +o, 1711 oo
   Survival Rates       :   20 ++,   24 +o,   20 oo
   Reproductive Rates   :    6 ++,    6 +o,    0 oo
```

Prediction
(enter your own, have your instructor initial it):

Results:
What were the final allele and genotype frequencies?

	Genotype frequency ++	+o	oo	Frequency of + allele
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 succeding 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 one 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 for us 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 which will allow you to compare changes of allele frequencies in populations of different sizes. To limit population sizes, you should input appropriate values for K and post-k. 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?

Here 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 (i.e., 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 o allele, and began receiving immigrants from another population that contained a high frequency of + alleles. If the + allele produced a phenotype with no effect on survival or reproduction, but which dispersed readily, then the gene flow would favor that allele.

Set up you EVOLVE data as follows:

```
   Title                :   12. GENE FLOW
   Seed                 :   (choose your own 2-digit number)
   Initial Population   :      0 ++,      0 +o,   2000 oo
   Carrying Capacity    :   5000
   Post-crash Pop. Size :   3000
   Survival Rates       :     20 ++,  20 +o,  20 oo
   Reproductive Rates   :      6 ++,   6 +o,   6 oo
   # immigrating        :    500 ++,  50 +o,   0 oo
   % emigrating         :      0 ++,   0 +o,   0 oo
   Number of Generations:    109
```

Prediction :
(enter your own, have your instructor initial it):

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 oo individuals. How would the results of this experiment differ if the emigration rates were set to 30%, 30% and 0% for the ++, +o and oo genotypes respectively?

MUTATION

Mutation is of tremendous 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, mutational equilibrium, etc. 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 7.5%. 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, have your instructor initial it):

Experiment:
To simulate a population in which a mutation has just occurred, you will want an initial population with many WWs, 1 +o and 0 MMs. 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 using other random number seeds.

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 deletarious. What then is the fate of most such mutations? What about mutations that have no affect 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).

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?

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 population *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 possiblity, of course is that natural selections 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. Here is an experiment that illustrates this idea. 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.

Experiment:
Consider a population on an island where selection favors a recessive allele, o. Perhaps individuals carrying the dominant + allele have seeds with long hairs like milkweed that tend to be blown off the island, while oo 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 experiment like the following in which hairy seeds are blown onto the island from a mainland where selection favors mobile seeds:

Title	:	15. SEL VS GENE FLOW
Seed	:	(choose your own 2-digit number)

```
   Initial Population   :   1711 ++,   0278 +o,   0011    Survival Rates
:     20 ++,  24 +o,  24 oo
   Reproductive Rates   :      6 ++,   6 +o,   6 oo
   # immigrating        :    500 ++,  50 +o,   0 oo
   % emigrating         :      0 ++,   0 +o,   0 oo
   Number of Generations:    109
```

Results:
Describe the changes in allele and genotype frequencies in this population 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?
.page

Chapter 6:
INTERMEDIATE EXERCISES

This chapter provides some additional exercises to sharpen your understanding of evolution. They differ from those in the previous chapter by being more quantitative and more abstractly formulated; we will also examine the impact of selection on population sizes. 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

These initial exercises, like those in Chapter 5, 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 and study their impact on growth of population size. Use the following tables to answer quesions 1-5.

a.

"Phenotype"	Genotype ++	Genotype +o	Genotype oo
Surviv. rate	25%	25%	20%
Reprod. rate	3	3	3

b.

Genotype ++	Genotype +o	Genotype oo
50%	30%	30%
2	3	3

c.

'Phenotype"	Genotype ++	Genotype +o	Genotype oo
Surviv. rate	40%	40%	58%
Reprod. rate	1	2	2

d.

Genotype ++	Genotype +o	Genotype oo
12%	18%	24%
9	9	9

e.

'Phenotype"	Genotype ++	Genotype +o	Genotype oo
Surviv. rate	60%	75%	65%
Reprod. rate	2	1	1

f. None of above

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

a.

Genotype	Fitness Absolute	Fitness Relative
++		
+o		
oo		

b.

Fitness Absolute	Fitness Relative

c.

Genotype	Fitness Absolute	Fitness Relative
++		
+o		
oo		

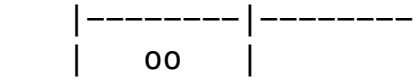
d.

Fitness Absolute	Fitness Relative

e.

Genotype	Fitness Absolute	Fitness Relative
++		
+o		

f. None of the above



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 whether the technique we use to study some phenomenon might obscure interesting points.

One such question might be whether it makes a difference whether 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?

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

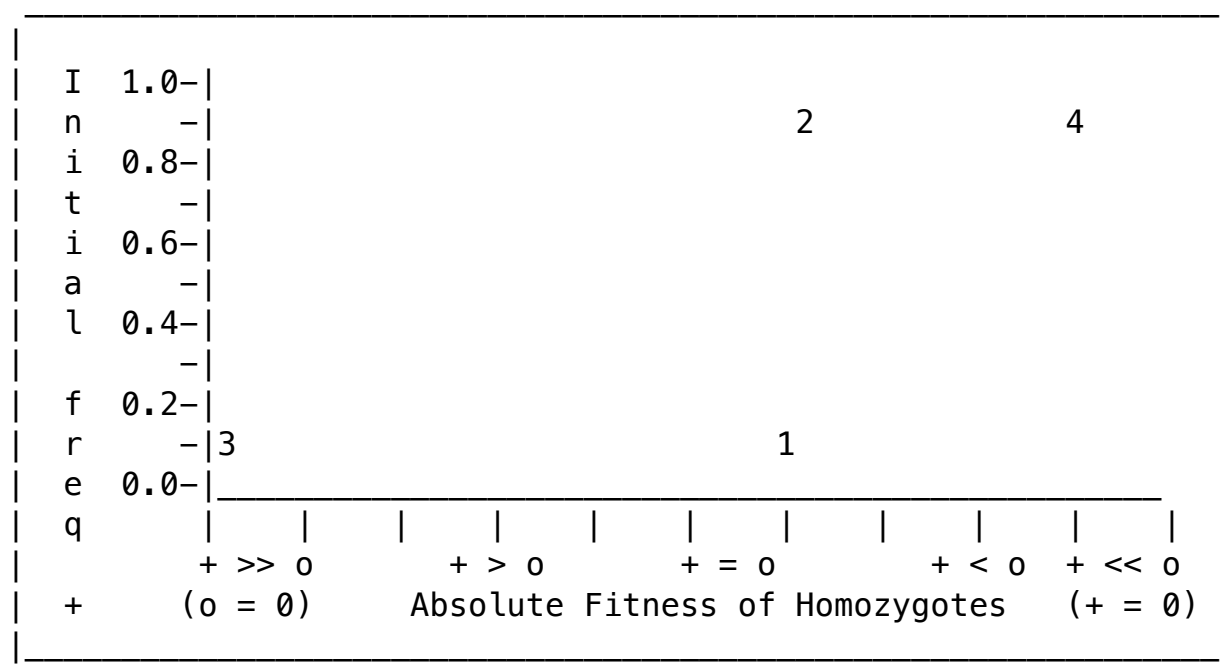
What do your results suggest about the usefulness of these coefficients in evolutionary biology? How do your results bear on the ecological concepts of r- and K-selected populations?

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 +o genotype is more fit than either of the homozygotes.

A convenient metaphore that may be helpful in planning experiments is to think of yourselves as exploring a landscape on the plane illustrated below. Four runs are outlined for you in the table below. You should make these runs and other runs of your own design.



Make several runs in which the fitnesses of + and o homozygotes are equal, but which start with different initial allele frequencies (e.g., the 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 (you needn't record it, but it is a useful exercise to test your understanding) 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 the surface before writing your answer to the question.

Run #	K	Post-K	Initial population ++	Initial population +o	Initial population oo	Initial frequency of + allele
1						0.05

Survival rates ++	Survival rates +o	Survival rates oo	Absolute fitness ++	Absolute fitness +o	Absolute fitness oo	Frequency of + allele in generation 125
50	50	50	1.5	2.0	1.5	

Reproductive rates ++	Reproductive rates +o	Reproductive rates oo	Relative fitness ++	Relative fitness +o	Relative fitness oo
3	4	3			

Run #	K	Post-K	Initial population ++	Initial population +o	Initial population oo	Initial frequency of + allele
2						0.95

Survival rates ++	Survival rates +o	Survival rates oo	Absolute fitness ++	Absolute fitness +o	Absolute fitness oo	Frequency of + allele in generation 125
			1.5	2.0	1.5	

Reproductive rates ++	Reproductive rates +o	Reproductive rates oo	Relative fitness ++	Relative fitness +o	Relative fitness oo

Run #	K	Post-K	Initial population ++	Initial population +o	Initial population oo	Initial frequency of + allele
3						0.05

Survival rates ++	Survival rates +o	Survival rates oo	Absolute fitness ++	Absolute fitness +o	Absolute fitness oo	Frequency of + allele in generation 125
			0.0	2.0	1.5	

Reproductive rates ++	Reproductive rates +o	Reproductive rates oo	Relative fitness ++	Relative fitness +o	Relative fitness oo

Run #	K	Post-K	Initial population ++	Initial population +o	Initial population oo	Initial frequency of + allele
4						0.95

Survival rates ++	Survival rates +o	Survival rates oo	Absolute fitness ++	Absolute fitness +o	Absolute fitness oo	Frequency of + allele in generation 125
			1.5	2.0	0.5	

Reproductive rates ++	Reproductive rates +o	Reproductive rates oo	Relative fitness ++	Relative fitness +o	Relative fitness oo

EXERCISE 18: Modelling the real world: Sickle-cell anemia and the scientific literature.

It is always a good idea to test your scientific models against the real world. Often data needed to do such a test is 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 witout 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 result and discuss the factors that might have affected EVOLVE's accuracy.

EXERCISE 19: Plotting / q vs q.

As you may know from an examination of texts on theoretical population genetics, mathematical models of selection frequently derive the / q, 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. Tabulate the allele frequency (q) of the allele that increases in frequency at intervals of a generation or two (use printed summaries if they are available, or you can read the data off the screen if you

use a piece of paper as a straight-edge). Subtract adjacent values of q to obtain / q; then do a scatter plot of / q vs q. Be sure to note the survival and reproductive rates of the experiment. If members of a class each do a separate experiment assigned by the instructor, you will soon have abundant material for discussion.

For example, ask yourself what values of q have the highest values of / q in experiments 7A and 8A. Can you see how the graphs of allele frequency vs time that EVOLVE produced relate to the plots of / q 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 / q?

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

EXERCISE 20: Plotting / q vs population size.

When you are studying drift, it is more useful to graph / q 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 the points from exercise 11. (Give some thought whether you should plot the size of the population in the first or last year, or whether 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 if care is taken to ensure that everyone uses different random number seeds and keeps the other inputs the same. Record the allele frequecy 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. .page

Chapter 7: ADVANCED EXERCISES

The exercises in this chapter are designed to help you develope 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 and often mathematically sophisticated, rather than experimental, and a complete understanding of some of these models 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 have probably spent quite a bit of time using EVOLVE, but you probably do not really understand how it can produce realistic (reasonable?) results in such a wide variety of evolutionary situations. What we shall do in this exercise is to 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, specifically, the approach used in EVOLVE. An understanding of the basic approach of this type of model is very important in more detailed analyses of its results.

The Problem

Before coming to class consider how ~you would model the evolutionary situation described below. Please feel free to discuss this assignment amongst yourselves 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. Bring your notes to class and be prepared to discuss your approach. In class you will set up a model and "run" it by hand. After getting results we will see how this model relates to EVOLVE. Please bring a calculator to class.

Restrict your thinking to a species which is diploid, has sexual reproduction with obligate outbreeding, which is monoecious (hermaphroditic), and has 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 o) at one locus. If it helps you visualize an abstract problem, you may think of the locus as being pleiotropic, with o alleles producing spines which protect the organism against predation, and, when homozygous, a somewhat shorter intromittent organ that reduces effectiveness of copulation. Consequently, oo individuals have a higher survival rate than ++ and +o individuals, but a lower reproductive rate. Heterozygous, +o, 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	0.20	5.10
+o	210	0.30	4.60

oo	45	0.30	4.10
Total	500	----	----

Do not try to formulate an equation for such a population, rather consider the basic population process of birth, 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 discuss your approaches, then provide a summary of one approach and let you calculate the results to see how your answers compare to the model's.

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."

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 (i.e., selection coefficient = 1 and relative fitness = 0), then the formula for change in gene frequency becomes:

/ q = -q^u2^d/(1+q)

You might calculate the values of / q 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 sobering 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 might be

instructive to break your class 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. Do all of the teams come up with the same ideas?

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

Chapter 8:
SETTING UP EVOLUTIONARY EXPERIMENTS

INTRODUCTION

This chapter contains reference material on each of the parameters you may input to EVOLVE. 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.

GENOTYPE VARIABLES MENU

Starting Population: These three numbers determine the composition and size of the starting adult population of generation 1. They must not total to more than 4000.

For example, suppose you wish to study the fate of a new mutant allele (o) in a population of 1000 ++ individuals. You should use 999, 1 and 0 as the initial numbers of ++, +o, and oo 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 and you should be able to do so. See "Calculation of fitness & seelection coeficients" in chapter 9 if you do not know how to do this.

Natural Selection: Survival and Reproductive Rates.

Survival rates may vary from 0% for a lethal genotype to 100 for a genotype in which 100% of the young survive to adulthood. Obviously the former is not unusual in nature, while the latter is very unlikely.

Reproductive rates may range from 0 offspring per parent for a sterile genotype to 10 offspring per parent for a fertile one.

These are very important variables which determine the pattern of inheritance, and of natural selection, and greatly influence rate of population growth. You absolutely ~must understand them thoroughly.

Pattern of inheritance -- Choice of appropriate values for these variables will allow you to simulate any pattern of inheritance possible for a single locus with 2 alleles (e.g., dominance and recessiveness, codominance, heterozygote superiority, etc.). 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 ++, +o and oo genotypes respectively, then the + allele would be dominant and the o allele would be recessive. (Note that dominance has nothing do do with which genotypes have the highest survival rates, only with the homozygote that the heterozygote resembles. In this case, the heterozygote "looks like" 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 -- incomplete dominance.

Population growth rate -- The relationship of 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 the population grows slowly. As a bad example, consider a run 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 and average of 1 young for the next generation, i.e., 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 would depend on the relative proportions of the various genotypes in the population as well as 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 ++, +o and oo genotypes respectively. The + allele would be recessive (because the heterozygote has the same characteristics as the oo 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		
	++	+o	oo
Survival Rate	20%	20%	14%
Reproductive Rate	5	6	7
Absolute Fitness	100%	120%	98%
Relative Fitness	0.83	1.00	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 reproduhat would happen to the population. The oo genotype, although it has a low survival rate, has a high reproductive rate. Would the o 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 oo 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) one can see that the ++ homozygotes have only 83% of the fitness of the heterozygotes, while the oo 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 Dandilion, and the seeds have plumes for wind dispersal. The alleles might be conceived as affecting the length of the plumes, hence 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.

CHANGE OTHER VARIABLES MENU

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.

Population Size: Carrying Capacity and Post-Crash Population.

These parameters along with rates of survival and reproduction (and gene flow, if any) determine the size of the population. The carrying capacity may range from 10 to 5000, while post-crash may range from 2 to 4000. Obviously, the carrying capacity must be larger than the post-crash size.

The carrying capacity (equivalent to the ecologist's ~K) is the maximum size 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 carrying capacity, 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 50%, you would expect the number of adults to be about 10,000, twice the largest possible carrying capacity. 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 K = 20 and post-crash = 10; provided the population does not become extinct it will remain within those limits. Setting K = 5000 and post-K = 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 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 carrying capacity of 5000.

Random Number Seed.

This is a whole number between 0 and 99. It is used by the computer to simulate random mating and the effects of weather and other "random" factors.

EVOLVE takes the "seed" and performs a series of mathematical operations on it to produce a "random" number. That number is used by EVOLVE in various ways (to randomly pick pairs of parents from the adult population, for example), and as the seed for the next "random" number.

For example, when the program reduces the number of young to obtain the number of adults, it takes the survival rate and 'seed' you input 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 have an average value equal to the value you input, but vary slightly. If you input 60%, the survival rates in 5 consecutive generations might be: 61.8%, 66.3%, 53.9%, 57.5% and 60.3%. If you used another seed the 5 survival rates would be different but would still average out to about 60%.

Different 'seeds' will produce different series of 'random' numbers and hence slightly different runs. The same seed always produces the same series of numbers and, other things being equal, the same run.

This may seem complicated, but it is a practical way of simulating the sort of variations which occur in nature. For example, weather might be mild for a couple of years, then harsh, in an unpredictable fashion, or predators that take advantage of one phenotype might be unusually abundant in one generation and unusually scare in another.

If you use the same seed in different experiments you can be sure any variations in output are due to the 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.

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 110. Also, there are times when you may want to change variable values on a specific schedule.

PROGRAM OUTPUT

Aside from prompts and messages to the screen, EVOLVE prints three types of information: Summaries of input data, summaries of individual generations, and various graphs. The graphs may be further subdivided into three subgroups. Each of these types of output is discussed briefly below.

Printed Summaries

Each time you ask for printed summaries, you may select the initial and last generations to print, and may ask for summaries at particular intervals. The printed output will indicate the input data and a table showing the generation number, the number of adults of each genotype, and the frequencies of genotypes and alleles. Summaries of first and last generations are always printed to give starting and endpoints of a run.

Although graphs are sufficient for most situations, the data are inevitably rounded as axes are scaled, so the graphs may not be as accurate as necessary to answer a particular question. Thus, users may request optional summaries of intermediate generations. These are useful in specific situations where knowledge of the exact composition of particular generations is needed. However, users should not ask for summaries indicriminantly, for they greatly increase the time it takes for a run, as well as the volume of output.

Graphs

The graphs which may be output by EVOLVE fall into three categories: 1) frequencies of alleles or genotypes in each generation, 2) number of individuals in the population or of each genotype, and 3) selection coefficients. In general, the most useful graph is of allele frequency, followed by population size, genotype frequencies, numbers of each genotype and, finally, the coefficient graphs.

All of the graphs have a horizontal axis which is time in generations. Graphs of numbers (type two graphs) have a vertical axis which is simply number of individuals, and bars are used to show the values in each generation. Graphs of frequencies and coefficients have vertical axes with various meanings and are scaled from 0.0 to 1.0. Dots are used to indicate the value of + alleles and

++ homozygotes; vertical dashes are used for o alleles and oo genotypes, and solid lines are used for +o heterozygotes. If an allele or genotype is present in small numbers, the appropriate symbol may print on the zero line because of rounding. If the datum is really zero, nothing is printed on the zero line.

Frequency graphs. These two graphs are the most useful, for they show the relative distribution of alleles or genotypes in the population. Evolution may be defined as "change in allele frequency." In that case, a graph of allele frequency is the single best summary of what happened to the population. The allele frequency graphs are always symmetrical about a horizontal line at 0.50 because the frequencies of the two alleles must always sum to 1.0 ~(p + ~q = 1). The genotype frequency curves are more involved, but generally follow the Hardy-Weinberg relationship: p^u2^d + 2pq + p^u2^d = 1. Although this relationship does not apply precisely when selection, drift and/or gene flow operate, the actual frequencies are often fairly close.

Size of population, genotype numbers. Each of these four graphs shows the number of individual organisms in each generation. The graph of population size is useful in any situation where the population is small or numbers fluctuate drastically. Graphs of individual genotypes are useful at the introductory level to show students what is happening to the population, and to show the Hardy-Weinberg relationship mentioned above. 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 Js in the J-shaped pattern of growth, and the spacing between them clearly illustrate that r, the ecologists' "intrinsic rate of natural increase" evolves over time and is a measure of the average absolute fitness of a population.

Fitness and selection coefficients. Users may ask for graphs of selection coefficients of the alleles and of the genotypes. these four graphs are perhaps most difficult to interpret. They are meaningful only in runs where selection alone is operating, and do not correspond to the theoretical, constant selection and fitness coefficients used in equations by population geneticists. Rather, they are empirically derived estimates of fitness and selection from one generation to the next and differ from those calculated on the basis of reproductive and survival rates input by the user. Interested readers should read pp. 50-52 in Wilson and Bossert (A Primer of Population Biology, 1971, Sinauer Assoc., Stamford, Conn.) for a discussion of how these coefficients are calculated. These graphs ~are of use to advanced students studying the nature of selection coefficients and problems of quantifying selection from actual data.
.page

Chapter 9:
THEORETICAL NOTES

The purpose of this chapter is to provide additional perspective on some of the theoretical issues that lie behind use of programs like EVOLVE by scientists. Many of the details, such as how to derive Hardy-Weinberg ratios, etc. 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 o 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. freqencies of the two alleles will not change; 2. the ratios of the frequencies of the three genotypes (++, +o and oo) will be p^u2^d : 2pq : q^u2^d respectively, and will sum 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^u2^d : 2pq : q^u2^d (++ : +o : oo). Thus the Hardy-Weinberg "law" states that populations will remain in genetic equilibrium, i.e., 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 (i.e., are all of the same age, sex, etc.); 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 (i.e., factors, processes or phenomena that change allele or genotype frequencies):

1. Individuals are not equivalent (e.g., 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 (e.g., meiotic drive, gametic mortality); 5. individuals differ in their ability to survive and/or reproduce ~(natural ~selection); and 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 2 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 sum to one
(p + q + r = 1),
and the frequencies of the genotypes would be
(p + q + r)^u2^d = 1^u2^d,
or
p^u2^d + q^u2^d + r^u2^d + 2pq + 2pr + 2qr = 1.

Importance

Disregarding 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. 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 (if allele or genotype frequencies deviate significantly from expected values) we know that one or more of the Hardy-Weinberg assumptions have been violated and can proceed to determine which evolutionary force(s) affect 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 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 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 manipulation 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 and it is difficult for students to get a feel for the way factors interact in populations. 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 embodies an intuitively simple, yet realistic model. Depending on the values of input parameters, it 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. (See the discussion of the Random Number Seed in chapter 8 for more on the role of random numbers.) 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, changing only random number seeds each time, to get a feel for chance variations.

CALCULATION OF FITNESS AND SELECTION COEFFICIENTS

Despite the importance of measuring actual fitness and selection in nature and 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 middle of the 1960's did we have a method that gave us 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 (e.g., 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 one might calculate fitness and intensity of selection from real-world data is a matter of considerable controversy and there is a large body of literature devoted to the 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 hyposthes. 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 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 = 0.156
---
check: sum = 1.001 (accurate within 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^u2^d, 2pq and q^u2^d.

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) = 0.458
---
check: sum = 1.000

b. Expected genotype frequencies in generation 2:

freq. of AA = p^u2^d = (0.542)^u2^d = 0.294
freq. of Aa = 2pq = 2(0.542 * 0.458) = 0.496
freq. of aa = q^u2^d = (0.458)^u2^d = 0.210
---
check: sum = 1.000

3. Calculate the absolute fitness of each genotype R(genotype^u as the observed frequency divided by the expected.

R(AA^u = 0.455 / 0.294 = 1.548
R(Aa^u = 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 must have a fitness of 1.0). Absolute fitness is the measure that tells us something about whether 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 (e.g., 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.

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.

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 thn of growth and crash illustrated in EVOLVE's population graphs does provide some measure of ecological reality as a consequence of the way you specify limits for genetic drift. In addition, the steepness of the Js and the spacing between them clearly illustrate that r, the ecologists' "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 to the dispersal of individuals 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, it is likely that one would expect a higher proportion of the population to emigrate.

Despite its extreme 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 mislead into thinking you really understand evolution. Should you then have to apply that intuition in real evolutionary studies, you may find that that intuition 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.
.page

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 modelling 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 Biolmetric 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. .page

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 +o and 50 oo individuals, the frequency of the + allele would be ((2 x 20) + 30)/(2 x (20 + 30 + 50)) = (40 + 30)/(2 x 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.

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^uA^dHb^uA^d individuals are normal; homozygous Hb^us^dHb^us^d have severe, usually fatal sickle-cell anemia. Heterozygous Hb^uA^dHb^us^d individuals have sickle-cell trait, a mild anemia.

Deme - A local population within which mating is random.

Directional Selection – Pattern of selection that changes the frequency of an allele in a constant direction, either toward 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 o is dominant, then oo and o+ individuals will have the same genotype, and ++ homozygotes would have the recessive phenotype.

Emigration – Movement of individual orgainisms 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 idividual 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^u2^d:2pq:q^u2^d. Hardy-Weinberg equilibrium requires that certain conditions be met (e.g., 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.

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 modelled by EVOLVE.

Mutation - The process or event that produces an inheritable change in a the 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; e.g.,

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, berhavioral or other properties of an organism the develop through the interaction of genes and their environment.

Pleiotropy – Condition where 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.

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.
.page