

KING'S COLLEGE

BIOL 113L Evolution and Diversity Lab

Lab Manual

Fall 2014

Contents

Laboratory Notebook Guidelines.....	3
Lab Report Guidelines.....	7
Descriptive Statistics Lab.....	9
Calculating means and standard deviations in continuous data:.....	10
Standard deviation, whelk shells:.....	11
Data entry and manipulation:.....	11
Means and standard deviations:.....	11
Making a bar graph with EXCEL.....	13
Making a scatterplot with EXCEL.....	14
Descriptive Statistics and Graphing Problem Set.....	16
Inducing Evolution in Bean Beetles.....	20
Understanding and Preventing Plagiarism.....	22
Population Genetics and the Hardy-Weinberg Principle.....	23
Hardy-Weinberg Problem Set.....	28
The Evolution of Caminalcules.....	29
Review of Phylogenetic Trees.....	36
Phylogenetic Systematics Lab: Evolution of Ceratopsian Dinosaurs.....	39
Ceratopsian Character Guide.....	44
Ceratopsian Character list.....	46
Global Map of Ceratopsian Discoveries.....	49
Berta & Sumich (1999) Chapter 2 – Marine Mammals: Evolutionary Biology	

Laboratory Notebook Guidelines

Adapted from: <http://www.ruf.rice.edu/~bioslabs/tools/notebook/notebook.html#prepare>

Choosing a notebook

Pads of tear-out graph paper or spiral bound notebooks without pre-numbered pages are not acceptable. It must be impossible to tear out a page without leaving evidence. It is safest to select something that is clearly labeled as a laboratory notebook. It is also helpful to choose a notebook with carbon copies, so that you can take those copies with you to study and write lab reports and organize presentations. Composition books are also acceptable.

Preparing the Notebook

Please use a pen for all entries (ball point is best), so that the marks will not smear nor will they be erasable.

Put your name, course number, semester, and laboratory section on the outside front cover of the record. Put that same information on the first page inside, or on the inside front cover. If your notebook does not include a pre-labeled table of contents section, then reserve the next several pages for a table of contents by labeling the top of each page as Table of Contents and numbering each page. If your notebook does not have prenumbered pages, you may wish to use lower case Roman numerals, as in a standard publication. Next, number the next several pages with Arabic numerals in sequence, and you are ready to begin recording data.

What to enter

Above all, it is critical that you enter all procedures and data directly into your notebook in a timely manner, that is, while you are conducting the actual work. Your entries must be sufficiently detailed so that you or someone else could conduct any procedure with only the notebook as a guide. Few students (and not that many researchers for that matter) record sufficiently detailed and organized information. The most logical organization of notebook entries is chronological. If a proper chronological record is kept and co-signed by a coworker or supervisor, it is a legally valid record. Such a record is necessary if you or your employer are to keep your rights to your discoveries.

The bare minimum entries, for each lab study, should include:

1. title of the lab study
2. objective(s)
3. introduction
4. detailed procedures and data
5. summary

We usually record a lot more information in a laboratory notebook than we would report in a research paper. For example, in a published article we don't report centrifuge type, rpm, rotor type, or which machine was used. However, if a procedure is unsuccessful you may want to check to see that you used the correct rpm or correct rotor. Perhaps the centrifuge itself was miscalibrated. You would need to know which machine you used. In a research paper one does not report which person performed which tasks, because such information is useless to a third party. However in the notebook it is important to note who was responsible for what procedure. Again, you may need such information to troubleshoot your experiments.

Making entries

Someone else may need to consult your notebook sometime, so please make your entries clear and legible.

Organization: Depending on how your notebook is designed you may choose whether or not to use the backs of pages. If you leave them blank, put a corner-to-corner line through them to void all blank spaces. Some people use the backs for rough calculations, then void remaining blank space. You might also decide to save space (and trees) and use both sides of each page. Obviously you cannot use both sides with notebooks that are designed to make duplicate copies.

Introductory information: When you make your first entries of the day, start by entering the date. When you start each new page of a notebook enter the date next to the page number. Each page should be numbered and dated consistently. Most of us use the upper right corner of each page for date and page number.

Write a title for each and every new set of entries. Distinct sets of entries should be separated by using informative headings and by leaving a single space or two between individual sets of entries. Specific information can be more readily located that way. For a new laboratory study, write down a very brief introduction to the study, and list the objectives. If you have a specific hypothesis, include it next. The object is to make completely clear what you intend to do and why you are doing it.

Procedures: Record everything you do in the lab, even if you are following a published procedure. For example, if you started by obtaining a quantity of tissue from an instructor, then write down that you obtained tissue, describe it, note how much, what condition, etc. How much you write down is up to you, but any relevant information should be there. For example, it doesn't matter much if you received a chunk of liver in a red ice bucket or a black one. However, it does matter that the material was on ice. If you change a protocol in any way or decide between alternative methods, then the correct information must be recorded in the notebook. For example, a protocol for tissue fractionation may recommend centrifugation at 9400 x g, but we may decide to use 12,000 x g in the lab. The correct g force must be noted.

If you make a mistake, put a line through the mistake and write the new information next to it. Never erase or obliterate an entry. When you finish a page, put a corner-to corner line through any blank parts that could still be used for data entry. Every bit of every page must be legible and filled, either with information or with a mark that voids the section.

Data: You should record data directly in your notebook, making tables to organize the data. Also include any graphs, charts, statistics, and other summaries of data. If data is gathered or organized in the form of a computer printout, you should cut and tape it into the notebook, but secure it well and initial and date overlapping the entry. Make sure everything is clearly labeled including column headings, axes and legends on graphs, and especially pertinent units. Each table or graph (whether hand drawn or computer generated) should have a caption clearly explaining what type of data is being shown in the summary (i.e. Results of spectrophotometry readings at 560 nm, or Diagram of the dissection of the respiratory system of a frog).

Summary: When you have finished a project, summarize what you have accomplished. You don't have to draw conclusions, just indicate what sort of data or observations you collected, samples you saved (and where and how you saved them), or any other relevant information that wraps up the study. For a continuing study keep the summary extremely brief. In fact, if the notes are well organized and it is obvious where the study left off, you need write nothing more than "To be continued..." Summaries help maintain continuity. They indicate where the work left off and how it might resume.

Doing two things at once?

What if you are conducting two long procedures at once, each with long waiting periods? For example, suppose you are conducting a protein assay and preparing a gel for your samples out in the laboratory. Back in the cell culture room, you are harvesting and processing tissues for primary culture. Both procedures involve waiting periods, yet you will complete both tasks by the end of the day.

Simply use your best judgment. You could divide each page into columns and keep your two records side-by-side. You might date two consecutive pages, keeping both records separately. In either case, when you leave the laboratory for the day cross out any unused parts of a page that precede the last entry.

Continuation pages

What if you need more than one page for a project? With continuing research, that will always be the case. Proper use of continuation notes makes it possible to follow your path through a long experiment or series of experiments without having to leaf through every page of your notebook.

For example, let's say you labeled some protein samples with the radioisotope S-35, ran a gel, and placed the gel in a film cassette in order to produce an autoradiograph. During the two days your film is in the freezer, you devote all of your time to a cloning project that is part of an unrelated study. After you put your film cassette in the freezer, simply write Continued, page __, then enter the date and title of your other project, and continue to record information.

When you resume work on the protein samples, enter the date, write Continued from page __, and enter your autoradiography results. This way, everything you do in the laboratory is recorded chronologically, yet someone interested in following your progress could start from the beginning and follow every procedure on just that one study, from start to finish.

Are things getting too sloppy?

Perhaps your data records are scattered throughout the notebook, and you would like to summarize them. Go ahead. You may re-enter tables or figures any time you wish to organize your work a bit better. To prevent confusion over duplication of data you may put a line through a table or figure you intend to re-draw, initial and date the change, and note the page on which the re-organized data can be found. Just don't obscure any of the original entry.

Repeated procedures

So far you have been advised to record each step you perform in the laboratory, regardless of whether the procedure is published somewhere. However, once you carry out a procedure, you can refer to that part of your notebook, and only note changes you make. For example, the first time you measure your beetles you should write down the exact scale, how you handled the beetles, how long you left them on ice, etc. The next time, just refer to the name of the procedure and the appropriate page(s) of your notebook.

Loose materials

Suppose you enter raw data into a computer and have a printout with 400 pieces of data. Or, suppose you generate a graph using a software program. These items should be attached to the notebook itself, preferably using tape on all corners, so that these items cannot be easily torn off or removed. Added materials should be initialed and dated overlapping the entry.

Table of contents

It is important to have a detailed table of contents (TOC) so you and your instructor can find entries easily. Record all entries in the table of contents as you go. For this course, use a separate page in the TOC for each lab exercise, and include multiple levels in a table of contents. This means that the title for each lab should be at the top of a TOC page, and hierarchical subheadings should follow for easy location of specific parts of a study, such as methods, sets of data, etc. The idea is to enable someone (such as your professor) to find anything quickly. List each set of entries with dates and page numbers. You and your instructor must be able to find any specific entry quickly, without flipping through pages.

Lab Report Guidelines

The formal lab report serves several purposes in an introductory biology lab. It serves as an evaluative tool to determine how well you performed your experiment, how much you understand about what happened during and after the experimentation process, and how well you can convey that information in an organized fashion. It also provides you the opportunity to develop your written communication skills and, perhaps most importantly, to learn how science is communicated.

Lab reports should be approximately 4-6 pages in length, double-spaced, excluding tables and figures. However, the quality and completeness of the content is more important than the number of words or pages. Keep the report clear and concise.

Biology Lab Report Formats

A lab report has a format that includes five main sections, modeled after a primary scientific research journal article. The sections are the **introduction, materials and methods, results, discussion, and references**. Your report must be divided into these five sections, using each as a section heading.

Title: (not a section *per se*, but should be included nonetheless) The title should be descriptive and specific. Instead of "Plant pigments," the title "The effects of light intensity on photosynthetic pigment concentrations in *Spathiphyllum wallisii*" is much more informative

Introduction: The introduction of a lab report states the purpose of the experiment. Here you will also provide some relevant background information on the topic that includes the basic biology and definitions of terms that may be unfamiliar. The background information should provide the reasoning (rationale) behind your hypotheses. The background of a good introduction will make reference to outside sources, preferably primary sources. The background information is followed by a summary of what you intend to do in the experiment, a clearly stated objective, and a specific hypothesis.

You may find it easier to write the introduction after you have completed the materials and methods, results, and discussion sections of your lab report.

Methods: This section of your lab report provides a written description of the materials used and the methods involved in performing your experiment. You should not include a list of materials you used, but organize this section in paragraphs that clearly describe what you did and how you did it. This section should be written in past tense; after all it already happened. **DO NOT copy the instructions from your lab handout; that would be plagiarism!**

How much detail? The general rule is that this section should be informative enough and clear enough that someone who has not done the experiment before but has a general knowledge of biology could replicate your methods. But the information you include should not be overly detailed, for example, you don't have to explain how to work a micropipettor or that you labeled your tubes with tape or marked your field plots with red flags as opposed to blue. If you are unsure if a piece of information should be included, ask yourself if that part of the methods could or did influence the data you collected; if so it should be included.

Results: The results section should show your data *and* state what you found. You should save any interpretation and discussion of the meaning of your results for the discussion section.

What you show is a summary of your data with graphs, tables, figures and any other illustrations of the data, and not a table of the raw data. If you choose to include the raw data (or your instructor requires its inclusion) it should be given in an appendix at the end of your report. Choose the format and layout that best shows the differences between your treatments, in a way that allows you to address your hypothesis. Make sure each graph, table, etc. is sufficiently labeled (axes, columns) and that units are shown where ever needed. Each graph or table needs a caption consisting of:

Figure 1 (or Table 1): With a description of what is displayed in the figure or table. For example, "Mean (+/- 1 standard deviation) of chlorophyll a and b concentrations for leaves of species x, y and z."

Be sure to also include a written statement of your results in paragraph format, noting any patterns or trends observed and making specific reference to the data. For example:

Chlorophyll a concentrations were greater than chlorophyll b concentrations in species x and y, while chlorophyll b concentrations were greater than chlorophyll a in species z (figure 1).

Discussion: This section is where you summarize the overall results and then interpret your results in a biological context and fully discuss the implications. Relate your results to the background you provided in the introduction. State if your hypothesis was supported (not proven) or not. If your results are not as expected, propose alternative hypotheses. Here is where you should also note any possible sources of error and consider how that error could have affected the results you collected and your ability to test your hypothesis. If there is anything about your experiment that you think could be improved upon, provide suggestions for doing so. If given the opportunity to continue the research, describe the next step including how you would test any alternative hypotheses. Most research leads to more questions than answers.

References: Provide references to all sources of information, including any books, articles, lab handouts, lab manuals, diagrams or pictures, etc., that you used when writing your report. References should be initially placed after the sentence in the report (Smith and Jones 1982) or (Biology p.432) and then listed in full at the end of the report using the following basic format:

Author(s). Date. Article Title. Journal Name. Volume (Issue): page numbers.

Author(s) Date. Book Title. Publisher: City. Page numbers.

For additional help with references, refer to the guides provided on the King's Library website (<http://kings.libguides.com/citation>). Please use Council of Science Editors (CSE) style, name-year system.

Websites are **not** primary reference sources and should be used only when the original source cannot be located and then only if you can verify that the source is reliable.

Final comments

Find and read a primary research article. Pay attention to the format and writing style to get a better idea of what each section should include and how the information is best presented. Get in the habit of writing scientifically now. There is a campus writing center available to you if you have difficulty writing.

Please spell check, grammar check and proofread before you hand in your written work!

Descriptive Statistics Lab

*** **Important note:** “Datum” is the singular form, “data” is the plural form!!! ***

Measure 2 continuously variable characteristics for each of 2 related groups of organisms. For this lab, measure the length of the index and ring fingers on the left hand (2 characteristics) in men and women (2 groups).

Finger length is a continuous variable because the value can be any real number. This is compared to a categorical variable with values that fit into distinct categories [nominal and ordinal as a subset of categorical??] such as brown or blue eyes. An ordinal variable, such as the order of your birth in your family, (e.g. you can be the first- or second-born etc., but you cannot be the 1.534th born)

You should take measurements on at least 8 individuals in each of the two groups (or as instructed by your professor). Record the data in your notebook using a table similar to this:

Subject	Sex	Index finger length (mm)	Ring finger length (mm)
1			
2			
Etc.			

Calculate a mean and standard deviation ($\bar{x} \pm SD$) for each variable in each group (see below for how to do these calculations). You should thus have 4 means and standard deviations.

Why we use standard deviation: Standard deviation gives us a quantitative means of comparing the variation among populations. For example, the populations below all have the same mean, but we wouldn't say that the populations are all the same. Which of the populations is most variable? Which is least variable? Take a guess.

	Tree heights in area 1 (m)	Tree heights in area 2 (m)	Tree heights in area 3 (m)	Tree heights in area 4 (m)
	5	3	1	1
	5	3	2	1
	5	3	3	1
	5	5	4	5
	5	5	5	5
	5	5	6	5
	5	7	7	9
	5	7	8	9
	5	7	9	9
Average Height (m)	5	5	5	5

Standard deviations give us a way to compare the populations:

	Tree heights in area 1 (m)	Tree heights in area 2 (m)	Tree heights in area 3 (m)	Tree heights in area 4 (m)
	5	3	1	1
	5	3	2	1
	5	3	3	1
	5	5	4	5
	5	5	5	5
	5	5	6	5
	5	7	7	9
	5	7	8	9
	5	7	9	9
Average Height (m)	5	5	5	5
Standard deviation (m)	0	1.63	2.58	3.27

Do the sizes of the standard deviations match your qualitative guess at variability?

Calculating means and standard deviations in continuous data:

Mean: $X = \frac{\sum x_i}{n}$ Standard deviation: $\sqrt{\frac{\sum (x_i - X)^2}{n - 1}}$

Example with sea shell lengths (mm):

n: total number of objects counted

x_i : value measured (small x) for object i

Mean, cone shells: $X = (10.1 + 11.5 + 9.6 + 13.2 + 8.2) / 5 = \mathbf{10.52 \text{ mm}}$

Mean, whelk shells: $X = (16.7 + 19.6 + 18.4 + 22.4 + 20.6) / 5 = \mathbf{19.54 \text{ mm}}$

Standard deviation, cone shells:

step 1:

$$\begin{aligned} \sum (x_i - X)^2 &= (10.1 - 10.52)^2 + (11.5 - 10.52)^2 + (9.6 - 10.52)^2 + (13.2 - 10.52)^2 \\ &\quad + (8.2 - 10.52)^2 \\ &= 0.1764 + 0.9604 + 0.8464 + 7.1824 + 5.3824 \\ &= 14.5480 \end{aligned}$$

step 2:

$$SD = \sqrt{\frac{\sum (x_i - X)^2}{n - 1}} = \sqrt{\frac{14.5480}{5 - 1}} = \mathbf{1.9071 \text{ mm}}$$

Standard deviation, whelk shells:

$$\begin{aligned}\sum (x_i - X)^2 &= (16.7 - 19.54)^2 + (19.6 - 19.54)^2 + (18.4 - 19.54)^2 + (22.4 - 19.54)^2 \\ &\quad + (20.6 - 19.54)^2 \\ &= 8.0656 + 0.0036 + 1.2996 + 8.1796 + 1.1236 \\ &= 18.6720\end{aligned}$$

$$SD = \sqrt{\frac{\sum (x_i - X)^2}{n-1}} = \sqrt{\frac{18.6720}{5-1}} = \mathbf{2.1605 \text{ mm}}$$

**** Means and standard deviations are often written as: $X \pm 1 \text{ SD}$ ****

Cone shells: 10.52 mm \pm 1.9071 mm

Whelk shells: 19.54 mm \pm 2.1605 mm

Data entry and manipulation:

1. Enter the data for one of the measured variables into an Excel worksheet. We'll use our example with data on sea shell sizes.
2. Type in column headings so it's clear what your data are

Sea shell number	Shell length (mm) cone shells	Shell length (mm) whelk shells
1	10.1	16.7
2	11.5	19.6
3	9.6	18.4
4	13.2	22.4
5	8.2	20.6

Means and standard deviations: Calculating these statistics by hand is useful to help you understand them, but we have much quicker ways. Below are 2 ways to calculate in Excel. Do it both ways. More commonly, scientists use statistic programs like Minitab.

Using the Data Analysis ToolPak:

1. Select "Data Analysis" in the Tools menu at the top of the page

[Data Analysis is an add-in for Excel. It should be to the far right of the Data toolbar. If it's not there, click the File tab. Click "Options" to open a dialog box. Click Add-Ins on the left-hand menu. At the top of the list of add-ins should be Analysis ToolPak. Click it then click "Go..." Check the Analysis ToolPak box, and click OK. Data Analysis should now be on the right end of the Data toolbar and ready to use.]

2. Select "Descriptive Statistics" in the window that appears

3. Put your cursor in the “Input data range” and click, then select the data you want to use to calculate a mean. As an example, to calculate the mean of the cone shells, select the numbers in the grey boxes:

Sea shell number	Shell length (mm) cone shells	Shell length (mm) whelk shells
1	10.1	16.7
2	11.5	19.6
3	9.6	18.4
4	13.2	22.4
5	8.2	20.6

4. On the Descriptive Statistics window, click on the button for Output Range, and in the box that appears to its right, click on any box on your Excel worksheet, as long as there are a lot of empty boxes below and to the right

5. Click on the box to the left of “Summary Statistics”

6. Hit “OK”

7. The output is shown below for our cone shell data:

Column1

Mean	10.52
Standard Error	0.852877482
Median	10.1
Mode	#N/A
Standard Deviation	1.907092027
Sample Variance	3.637
Kurtosis	-0.22970087
Skewness	0.411499631
Range	5
Minimum	8.2
Maximum	13.2
Sum	52.6
Count	5

8. Make a new chart with your means and standard deviations:
(ignore the grey boxes for the moment)

	Shell length (mm) cone shells	Shell length (mm) whelk shells
mean	10.5200	19.5400
SD	1.9071	2.1606

Using simple formulas in Excel:

Excel will quickly complete calculations for you if you know a few simple rules. Most importantly, Excel only knows you're entering a formula if you type an equal sign as the first character in a cell.

Formulas can be:

- as simple as “=3+4” where the cell will read 7 and the formula bar will read the formula as typed
- or something like “=B4+B5” where Excel will add the values in cells B4 & B5
- You can do the same thing by using the SUM formula : “=SUM(B4:B5)”

Two formulas that are useful for this lab are:

- =AVERAGE(insert data range) will calculate the mean
- =STDEV(insert data range) will calculate the standard deviation

Calculate the mean and standard deviations using these formulas and compare the results to those using the Data Analysis ToolPak.

One of the most important reasons for using formulas (& why you should also be very careful with them & your data) is that formula values automatically update when data are changed within the data range used in the formula. Try this – change one of the values in your dataset (For example, change 80.0 mm to 800 mm). The mean and standard deviation should automatically change. Change the value back to the original, correct value.

IMPORTANT: Below (and above) are instructions for making both bar and scatter plot graphs based on example data. You are to use these instructions to help you make graphs for the data that you collect now and in any future labs. You will need to know these skills for the rest of your career, so make sure you understand how to make and label a graph and when it is appropriate to use a scatter plot vs. a bar graph. We start with instructions using the cone and whelk shell data so you can practice.

Making a bar graph with EXCEL (Instructions are for Excel 2010 – use an online tutorial or):

1. Select the values for the means (grey boxes in our example above)
2. In the Insert menu, click to create a 2D Column chart. Your chart will appear within your excel window.
3. Right click on the “Series” legend next to the bars in your graph. Chose “select data”, then click “edit” under the “Legend Entries (Series)”. A box appears. Type “Shells” as the Series Name, and click OK. Now click “Edit” under the horizontal axis labels. Now click in the “Axis” labels range and then back in the excel spreadsheet drag your cursor over the “cone shells” and “whelk shells” column labels, and labels will appear on your graph. Then Click OK.
4. For your dataset on finger lengths in men and women, a series legend is useful. However, notice that it's not useful if you're simple graphing average shell length of cone and whelk shells.
5. Now click back in the white area to the right of the graph. Then click the “layout” tab under chart tools, and the “chart title” tab, choose above chart, and name your chart in the text box above the graph. Something like “Shell lengths compared” would be appropriate for the example above.

- Use the same procedure as above but choose “axis titles” and label the “primary horizontal” and “primary vertical” axes “Shell Types” and “Average Length” respectively. Choose “Title Below” axis for the horizontal and “Rotated Title” for the vertical
- There, your graph is constructed and appropriately labeled so that others can understand what you are comparing and why. Now on to error bars.

Adding error bars to a graph (same procedure for any graph type)

- In the Analysis section of the Layout toolbar (you may need to click on the chart box to see the Layout toolbar within Chart Tools), click Error Bars and “**More Error Bars Options...**” This is important – clicking “Error Bars with Standard Deviation” will give you meaningless error bars (Excel should not give you this option).
- In the window choose the “Both” button to display both plus and minus 1 standard deviation. Then Click “Custom” at the very bottom and Click “Specify Value”
- Click in the “Positive Error Value” and highlight your standard deviation values for cone and whelk shells in your excel chart. Do this same procedure for the negative Error Values and then click OK.
- Now you have both negative and positive error bars in your graph that represent standard deviation above and below the mean.

Note: You can change just about anything (size of the chart, font on the axes labels, color of the bars, the background, etc.) by clicking on parts of the chart. Try this out and play around with changing things to make the graph look the way you’d like it to.

Making a scatterplot with EXCEL: [cone and whelk shells don’t lend themselves to this type of chart]

- Enter your data in columns with your labels in the top row and your x-axis values (the independent variable in column A). Your y-axis (dependant) values should be in columns B, etc.

Men		Women	
Index finger (mm)	Ring finger (mm)	Index finger (mm)	Ring finger (mm)
7	12.3	6.6	11.4
8.1	13.2	7.9	12.4
7.5	11.5	6.9	11.8

- Click on the insert tab and Insert a simple X Y (Scatter) plot (that hs only data point markers without lines that connect dots or show trends). You will get a blank graph; not worries, we will fill it in.
- Right click in the chart area and then choose select data. Under the “Legend Entries” click add. Under “Series name” highlight “Men” back in your excel chart, this will name the series “Men.” Click in the “Series X values” Then highlight the index finger length column of values for men. Click in the “Series Y values” and highlight the ring finger length column of data for men. Click “OK”. Now you should see three points on your graph for the “Men” data set.
- Now add a “women” series to your graph by repeating the same procedure above. Then click OK.
- Fill in the chart title, and the labels for your x- and y-axes using the skills you learned from the instructions in bar graph section above.

6. You can change many aspects of how the graph looks at this point by double clicking on the axes, legend, etc. The range shown on the x and y axes can be particularly important; when might you want an axis to start at zero, and when might you want to use a different number?

Include in your notebook all the basic parts of a notebook entry. Be sure to include:

Statistics:

1. Calculations of the means and standard deviations for both variables for your 2 groups. **SHOW YOUR WORK!** [These wouldn't typically be calculated by hand in a scientific notebook, but doing it here will help you better understand standard deviation.]

Graphs:

1. A bar chart showing the means and standard deviations of one of your variables for the two groups. Be sure the chart is labeled in a way that adequately describes what the chart is showing.
2. An X-Y graph that plots the two measured variables against each other **by group**. See this course's notebook guidelines for the proper way to attach printed pages, and be sure the chart is labeled in a way that adequately describes what the chart is showing.

Descriptive Statistics and Graphing Problem Set

Due:

Complete this exercise and answer the questions about graphing data and the use of descriptive statistics. Some are more difficult than others, but with some thought, you should be able to answer all of them. If you have trouble with this exercise, it would behoove you to see your professor during office hours as something much like this may appear on a quiz.

1. Does the amount of fertilizer I use make a difference in the amount of corn produced? Make a graph that answers the question. What does it tell you?

Fertilizer volume (ml/plant)	Weight of corn produced per plant (g)
0	357
1	348
2	470
2	358
3	468
3	491
4	502
4	515
5	622
5	588
6	600
6	630
7	622
7	634
8	641

2. I counted the number of flowers produced by two varieties of pea plants. One variety produces red flowers, and the other produces white flowers. I want to know if there is a difference in the number of flowers produced by each variety. How would I graphically depict my results? What are the variables on each axis?

3. There are two populations of a species of polychaete worms I am interested in. One population lives in sediment contaminated by the metal Cadmium (Cd). The other population lives in unpolluted sediment. Cadmium is toxic to worms. In the lab, I measured the survival of worms taken from each location. I had 20 buckets of worms from the polluted habitat and 20 buckets of worms from the unpolluted habitat. Half of the buckets of worms from both the polluted and unpolluted habitats were raised in the presence of Cadmium and the other half were raised without Cadmium. I measured survival in each treatment. The survival data are listed below:

Average Survival	Original Habitat	
	Polluted	Unpolluted
With Cadmium	71%	11%
Std Dev	+/- 3%	+/- 5%

Without Cadmium	88%	96%
Std Dev	+/- 4%	+/- 2%

What kind of graph should you make to compare different groups?

Which of these groups do you want to compare to each other? Why? (Hint: there is more than one interesting comparison here. Justify your choice of the most interesting comparison).

4. Characteristics of the small aquatic invertebrate *Daphnia melanica* differ among populations. What does the following graph tell you about *Daphnia* biology? [Ignore the asterisks.]

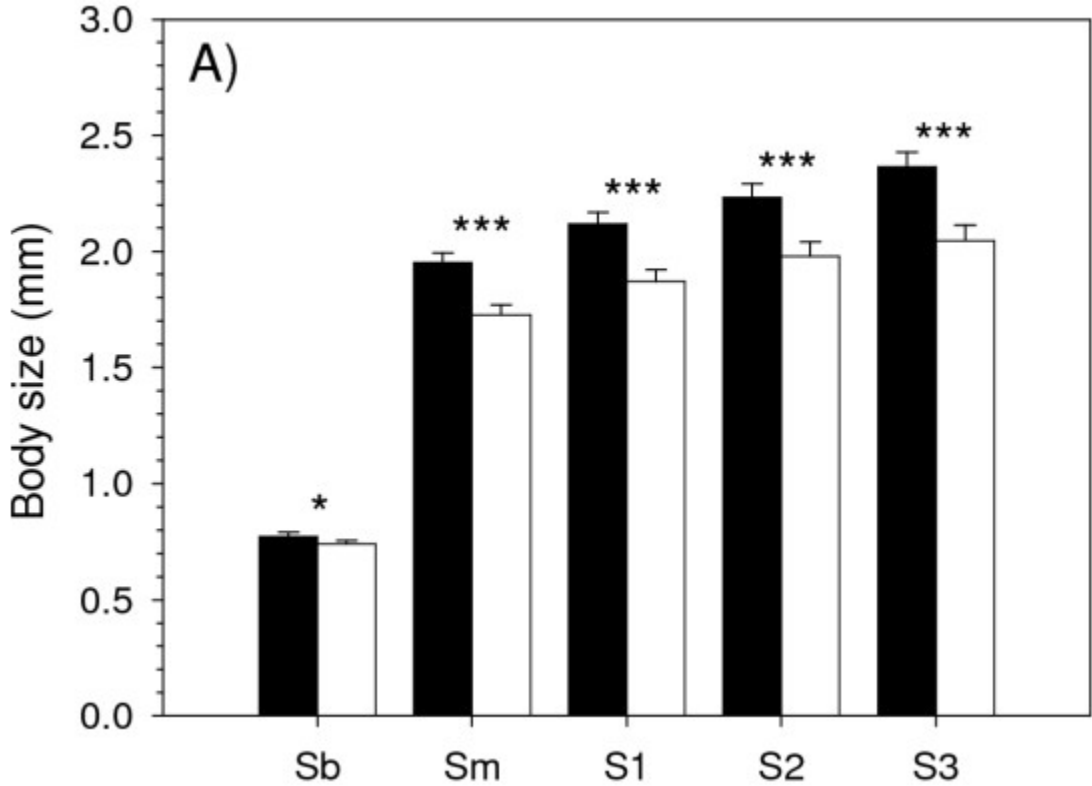


Figure 1: Body size of *Daphnia melanica* at different life stages in lakes with and without fish. Black bars represent populations of *D. melanica* from lakes without fish, and white bars represent populations of *D. melanica* from lakes with fish. The labels on the x-axis represent different life stages of *D. melanica* with the youngest stage on the left, and age increases to the right).

5. The body depth of a species of fish that lives in the Colorado River varies with the flow rate of the river. Individuals that live in fast moving water have shallow bodies (i.e. a low body depth) and those in slow moving water have deep bodies (i.e. a high body depth). How might you depict this relationship graphically?

6. A scientific study measured the hindlimb lengths of lizards and the diameter of the branches they perched on. The authors produced the following graph from their data. What does it show?

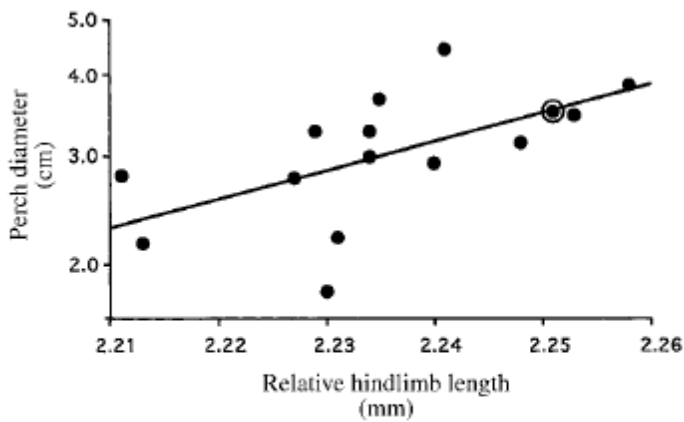


Figure 3. Relationship between perch diameter and relative hindlimb length. Perch diameter on a natural log scale. The circled point is Staniel. From Losos et al. (1997).

7. Examine the following two graphs. What does the first tell you about the groups measured in this experiment? What does the second tell you about the groups measured in this experiment? Why is there a difference in the interpretation of the two?



Figure 1. The mean percent (+/- standard deviation) of fish in silver salmon and king salmon diets.

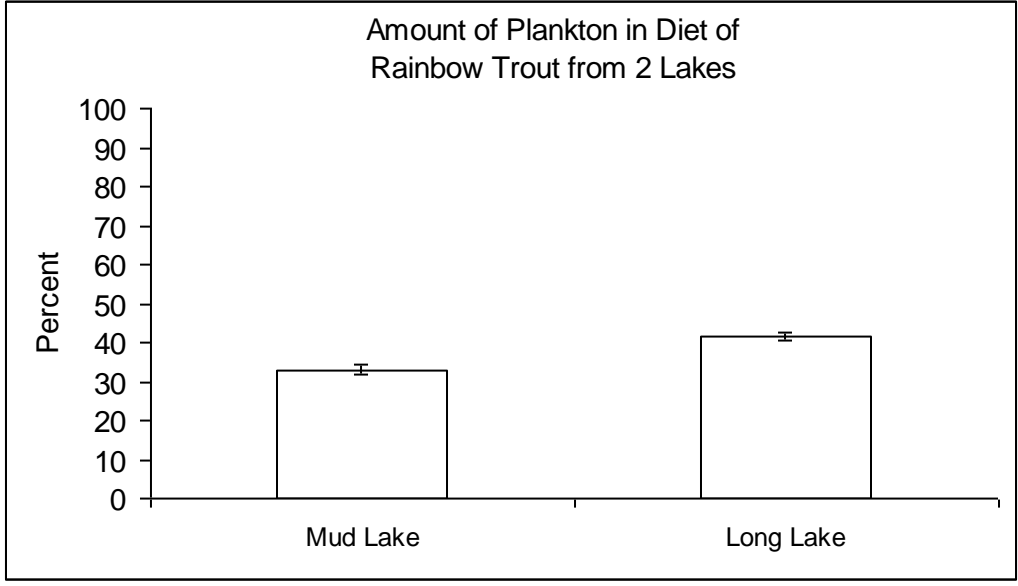


Figure 2. The mean percent (+/- standard deviation) of plankton in the diets of rainbow trout from Mud Lake and Long Lake.

Inducing Evolution in Bean Beetles

Adapted from:

Blumer LS and Beck CW. 2010. Inducing Evolution in Bean Beetles. Page(s) 25-35, in *Tested Studies for Laboratory Teaching*, Volume 31 (K.L. Clase, Editor). Proceedings of the 31st Workshop/Conference of the Association for Biological Laboratory Education (ABLE), 534 pages.

Objectives:

1. Design and perform an experiment to induce evolution by directional selection in bean beetles (*Callosobruchus maculatus*).
2. Evaluate control and experimental populations to measure evolutionary change.
3. Strengthen your laboratory and analytical skills and familiarity with the process of science.

Introduction

Evolution is defined “as a change in the genetic composition of a population from generation to generation” . Although natural selection is typically the most potent cause for evolution and is the principal cause for evolutionary change, other processes, such as mutation, migration, and genetic drift also can cause evolution (Freeman and Herron 2007).

Natural selection occurs when a trait, such as body size, varies in a population and individuals differ in their survival and reproductive success as a consequence of the particular character of a trait. For example, if adult body mass varied in a population and the risk of predation were greater among the smallest individuals in the population, then the larger individuals would have greater survival and consequently greater reproductive success than the smaller individuals. Directional selection, such as that described for body mass, may result in directional evolution if the variation in the trait (body mass variation in this example) were caused by genetic differences among individuals. In other words, if the variation in body mass were heritable, then directional selection on body mass would cause directional evolution in body mass.

The other potential causes for evolution (mutation, migration, and genetic drift) are real, but all cause random phenotypic changes in a population. Mutation is the spontaneous change in the genotype of an individual that may cause a change in the phenotype of the offspring of that individual. Migration is the movement of individuals into a population (immigration) or movement out of a population (emigration). In either type of migration, the mean value for a trait and variation in that trait in a given population may change as a result of such movement. Genetic drift results from random genotypic change that can occur when a population has very few individuals among whom reproduction occurs. When a population contains few individuals, even random mating may result in the loss of alleles and an increased frequency of homozygous genotypes compared to populations with greater numbers (Futuyma 1986). Genetic drift is a form of reproductive sampling error. To the extent that random changes in genotype frequencies result in changes in phenotypes, phenotypic evolution may occur as a consequence of genetic drift.

In this study, you will design and conduct experiments to induce evolutionary change in an insect species, bean beetles (*Callosobruchus maculatus*). Bean beetles are agricultural pest insects of Africa and Asia. Females lay their eggs on the surface of beans (Family Fabaceae). Eggs are deposited (=oviposition) singly and several days after oviposition, a beetle larva (maggot) burrows into the bean. At 30°C, pupation and emergence of an adult beetle occurs 25-30 days after an egg was deposited. Adults are mature 24 - 36 hours after emergence and they do not need to feed. Adults may live for 7-10 days during which time mating and oviposition occurs (Mitchell 1975). Adult body mass, linear body dimensions, and egg-to-adult development-time are variable traits in *C. maculatus*. Consequently, these easily measured traits are candidates for inducing evolutionary change in laboratory populations.

Experimental Design

Address the following questions. Come to class ready to discuss your answers.

1. What is the importance of trait variation if you were to induce natural selection?

2. Design an experiment or set of experiments to test the hypothesis that you can induce evolution in a trait by natural selection.

3. What assumption must you make about the cause for trait variation if selection were to result in evolution? Can natural selection result in evolution in a trait if variation were caused entirely by environmental variation?

For each of the experiments you designed above, you should:

a. Predict the possible outcomes for the experiment that would support your hypothesis

b. Identify and list the variables you would manipulate in your experiment

c. Identify and list the variables you would keep constant in your experiment

d. List the data you would collect to determine if your predictions were true

Literature Cited

Freeman, S. and J.C. Herron. 2007. *Evolutionary Analysis*. 4th edition. Benjamin Cummings. 800 pages.

Mitchell, R. 1975. The evolution of oviposition tactics in the bean weevil, *Callosobruchus maculatus* F. *Ecology* 56:696-702.

Futuyma, D.J. 1986. *Evolutionary Biology*. 2nd edition. Sinauer Associates. 600 pages.

Understanding and Preventing Plagiarism

Access the Purdue OWL resource on plagiarism <<http://owl.english.purdue.edu/owl/resource/589/01/>>

Read the section, "Avoiding Plagiarism," in "A Short Guide to Writing About Biology," pages 13-15.

Answer the following questions.

- 1) T/F Presenting someone's work as your own is clearly plagiarism, even if you change some of the words and rearrange the sentences.
- 2) T/F When you write about someone else's written or spoken ideas as long as you don't use the same wording they used, citations and references are not necessary.
- 3) Plagiarism can be intentional or unintentional. Give examples of both, and situations or circumstances in which they could occur.
- 4) Some forms of plagiarism are unquestionable. Others are less obvious (in fact, you may have previously been unaware that they were plagiarism). Give examples of both kinds.
- 5) Describe (obviously in your own words!) the difference between the things that require citation and those that don't.
- 6) What does "common knowledge" mean? What are some tips for deciding if something is common knowledge?
- 7) The Purdue OWL website describes "Best Practices." Which do you already practice? Which do you feel would most helpful to implement and why? What suggestions are given in Pechenik to avoid plagiarism?
- 8) What is the plagiarism policy in this course? Other courses you are taking at King's?
- 9) What are the possible consequences of committing plagiarism at King's?
- 10) Give your answers and the reasoning behind them to questions 1-8 in "Safe Practices: An Exercise" from the Purdue OWL website.

Please submit this assignment on Moodle prior to next week's class.

Population Genetics and the Hardy-Weinberg Principle

Microevolution is defined as a change in the frequency of different alleles at a gene (locus) between generations. In other words, if in one generation the frequency of A = 0.5 and a = 0.5 (or $p = q = 0.5$), and in the next generation the frequencies have changed to A = 0.6 and a = 0.4, then evolution has occurred (*remember*: $p + q = 1$). Individuals in a population either live to reproduce or they don't. If an organism doesn't reproduce, its genes are removed from the gene pool of the population. Therefore, evolutionary change is measured and detected at the population level. As a consequence, the study of evolution is quantitative and requires some sort of mathematical treatment.

As long as certain assumptions are not violated genotype frequencies will not change between generations and there won't be evolution. These Hardy-Weinberg assumptions are: no mutations, random mating must occur, no natural selection, large population size, no gene flow (ex. migration). The Hardy-Weinberg principle provides a simple analytical method to measure and predict gene frequencies. When one says that a population is in Hardy-Weinberg equilibrium, then the genotype frequencies have not changed. However, if one of the assumptions is violated then genotype and allele frequencies may change leading to evolution.

The objective of today's lab is to become familiar and comfortable with the use of Hardy-Weinberg principle. The HW equation will be used to calculate allele and genotype frequencies from phenotype frequencies obtained by students in this lab. We will also test the Hardy-Weinberg theorem by carrying out a series of random 'matings'.

The Hardy-Weinberg equation is: $p^2+2pq+q^2=1$.

This equation states that given the knowledge of the frequency of 2 alleles, we can calculate the frequencies of the genotypes and that the frequency of all of the individuals in a population add up to 1.0 (or 100%). This means that in a simple two allele system, one allele is designated p and the other q . If that system is one in which one of the two alleles are a dominant and the other a recessive, then p^2 represents the frequency of individuals with the homozygous dominant genotype, $2pq$ represents the frequency of individuals with the heterozygous genotype, and q^2 represents the frequency of individuals with the homozygous recessive genotype.

In today's lab we will sample phenotype frequencies for 1 character in peas and 2 characters in humans that are genetically controlled. We will also model allele change across generations with pop beads. We will also use data that was previously collected on cat coat color and use the data to become familiar with working with the Hardy-Weinberg equations and to test the Hardy-Weinberg theorem:

- 1) Pea color
- 2) Pop beads
- 3) Ability to taste phenylthiocarbamide (PTC) and sodium benzoate
- 4) Piebald spotting of cats

A. Seed color in peas, *Pisium sativum*

Each group has a population of peas. Determine the phenotype frequencies of your population. The yellow pea allele is dominant, green is recessive (just like in Mendel's studies).

Classwork 1: Assume your population is in Hardy-Weinberg equilibrium and calculate allele and genotype frequencies. (We will work through this as a class.)

B. Allele frequencies across generations

In this exercise we will be using pop beads to model a certain gene with two alleles, yellow and red. You will receive your initial genotype. Now do the following:

- Write down your initial parental (P) generation genotype.

- To ensure random mating, be completely uninhibited and choose anyone in the class (male or female) as your mate. NO REFUSALS!
- Each couple will have two offspring. Each parent contributes one allele to each offspring. If you are homozygous, you already know what allele you'll contribute. If you are heterozygous, determine your contribution by flipping a coin.
- The parents die after reproduction, so you will take on the genotype of one of the two offspring. (The other parent takes on the identity of the other offspring.) Record your new F₁ genotype.
- Chose a partner at random and follow exactly the same procedure, recording your new genotype each time. Carry this out for 5 generations.

Classwork 2: Determine the genotype and allele frequencies for the class using the P and F₅ generations.

C. Testing your ability to taste sodium benzoate and PTC

You will divide up into groups of about 3-4 students and obtain an envelope labeled A, an envelope labeled B, and an envelope labeled C. These envelopes contain taste test papers. PTC typically tastes bitter, while sodium benzoate tastes sweetish-bitter. The ability to taste PTC is dominant and the ability to taste sodium benzoate is also dominant. Therefore, it is likely that some of you will not taste anything (recessive). Each member of the group should taste one strip from each envelope. You should keep track of the number of tasters and non-tasters in your group. When finished, you will add your group's data to the class data on the board or in Excel.

When every group is finished, the contents of each envelope will be revealed.

Classwork 3: Again assuming that the population is in H-W equilibrium calculate allele and genotype frequencies for both characters using data collected by the class.

D. Using cat coat color to test the Hardy-Weinberg Theorem

Consider the following data collected on cat piebald spotting or the proportion of white in the cat's coat. The inheritance of this character is an example of incomplete dominance. No white fur is the homozygous recessive trait (ss), greater than 50% white is homozygous dominant trait (SS) (all white is a different trait that we won't consider here), and less than 50% white is the heterozygous trait (Ss).

Phenotype	Genotype	Number of cats
> 50% white	SS	18
< 50% white	Ss	104
No white	ss	109

Table 1. Data on piebald spotting in cats (Christensen 2000)

Classwork 4: Calculate allele frequencies in this population of cats. Calculate the *expected* genotype frequencies under H-W equilibrium. Are the actual (observed) and calculated (expected) frequencies different? If so, by how much, and in which direction for each genotype? Do you think these differences are meaningful and indicate evolution in this cat population?

E. Testing Hypotheses

When conducting a scientific investigation we usually arrive at a conclusion by either refuting or supporting hypotheses. Without doing further analysis you cannot be certain that the differences in frequency between the observed number of cats of each phenotype and those you calculated as expected under H-W equilibrium are meaningful, or could have happened by random chance alone. (Do not confuse this with the affect that chance has in small populations, genetic drift.) During our hypothesis testing when we say "by chance alone" what we

mean is: could you still see these differences by chance, BUT while still under the assumptions of H-W equilibrium.

In our present case we wish to test the hypothesis that our cat population is different from the expectations of Hardy-Weinberg equilibrium. However, first we must construct our hypotheses. It is customary to state your hypotheses in the form of a null hypothesis and an alternate. We usually designate the null hypothesis as " H_0 " and the alternate as " H_a ". The null hypothesis is often a statement of no difference, hence the word "null", meaning that the difference between the two conditions is of no consequence, effect, or the difference is insignificant. The alternative is obviously a statement of difference. For example:

H_0 : There is no significant difference between the observed frequency and that expected by H-W equilibrium.

H_a : There is a significant difference between the observed frequency and that expected by H-W equilibrium.

We then seek to reject a hypothesis and in doing so we might support the other. In this case if we find a significant difference we have grounds to reject the null and accept the alternate hypothesis. So how do we do that? Let's find out.

What would really help us here is a number. What we want to know is the probability that the results we have found could have occurred by random chance. Say for instance, that there is only a 5% probability that the differences we obtained between observed population and the expected (H-W) could have occurred by random chance alone. Well, that is a pretty small probability, and gives us confidence that the results we have found are due to real differences and not random chance. If the probability was 80% that they occurred by random chance then you would not be able to reject your null hypothesis. Why? Because it is entirely possible that the differences you are seeing are not really meaningful.

So... How do we get these numbers in the first place? We use a type of statistical test that is appropriate for answering our question. In this case the test is called a Chi-square (X^2) Test. There are many other types of tests but we won't worry about that right now. ☺

As an example, let's imagine that in a population of 30 cats the genotypes are distributed as SS = 4, Ss = 23, and ss = 3. From these we can get the individual allele frequencies. They would be S(p) = 0.52 and s(q) = 0.48. We plug these values into the H-W equation and obtain the expected genotype frequencies under no evolution. Thus we obtain SS = 8, Ss = 15, and ss = 7 as our expected frequencies. Indeed, these do look different from the observed, but...are the differences significantly different? In order to answer that question we must calculate the Chi-square (X^2) test statistic using the following equation.

$$X^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$$

From our example we get:

$$X^2 = \frac{(4 - 8)^2}{8} + \frac{(23 - 15)^2}{15} + \frac{(3 - 7)^2}{7} = 8.55$$

Now we must determine the probability of getting a Chi-square value of this magnitude based on some cutoff value of significance. In some cases an investigator may use a 1% probability and in others a 5%. We will use a 5% probability or $P = 0.05$. What we want to know is the X^2 value that is associated with this 5% probability and whether our calculated value is larger or smaller than it. To do that we need two more things, a number called degrees of freedom (df) and a table that shows the critical cutoff values of the Chi-square distribution (Table 2).

Degrees of freedom is calculated as the number of classes minus the number of independent values we used in determining the expected genotypes ($df = c - iv$). The number of classes is three, because that is the number of different genotypes or different groups. We also used two values that are independent from each other to calculate the expected number of cats of each genotype. One of these independent values is the total number of students and the other is the allele frequencies. P and q are not independent of each other because $p+q = 1$, so we count the allele frequencies as one independent variable. Remember that in order to find the actual number of students of a given genotype you had to first use the allele frequency (that is one independent

variable) to find the genotype frequencies, and then multiply the genotype frequencies by the total number of students (the other variable). So $df = 3 - 2 = 1$. Armed with this we can now use the table of critical values. Degrees of freedom is in the first column and probability is written across the top row, remember 5% is also $P = 0.05$. Reading the intersection of these values on the table gives us a critical X^2 value of 3.84.

OK, now what? Well...what you are saying with this critical value is that there is only a 5% probability of getting a X^2 value ≥ 3.84 by random chance alone. Because our X^2 value actually exceeds the 5% cutoff ($8.55 > 3.84$) this means that there is a much smaller chance (one less than 5%) that we would obtain our value (8.55) by pure dumb luck and not a real observed difference. What we have determined is that our differences between observed and expected are significant based on our 5% cutoff.

Where do we go from here, back where it all started, our hypotheses. We now have grounds to reject our null hypothesis of no difference due to chance and we have supported our alternate that is due to real difference. We say that our value is significant at the $P < 0.05$. We also feel much more confident about our assertions because we can put a probability value on their occurrence rather than saying there is a difference, and leaving it at that. How many politicians back up their claims with real numbers? Science for the win!

Classwork 5: Using the observed and expected values that you calculated in classwork question 4 do a Chi-square test to determine if your differences are significant or not.

Questions:

- 1) Why were we able to test for Hardy-Weinberg equilibrium in the cat population but assumed H-W equilibrium in pea color and human tasting?
- 2) The yellow pea phenotype could be caused by which genotypes? The green pea phenotype? How could you determine the frequency of the heterozygotes?
- 3) Tasters of PTC and sodium benzoate have two possible genotypes. What are they? What are some possible ways to determine the genotypes of tasters?
- 4) Based on your hypothesis testing (classwork 5), what are your conclusions about evolution of piebald spotting in the cat population? What conditions could lead to the evolution of this character?
- 5) A different character in the same cat population is the presence or absence of all white coloration. Twenty-six cats were all white and 239 cats were less than 100% white (Christensen 2000). Can you tell which trait is dominant, based on the information provided?
- 6) Refer back to table 1. Which allele is dominant? Which allele is most common? Why are these terms not interchangeable? Does this change your answer to the previous question?

Table 2 Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of χ^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11	3.053	4.575	5.578	7.584	10.341	13.70	17.28	19.68	24.72
12	3.571	5.226	6.304	8.438	11.340	14.85	18.55	21.03	26.22
13	4.107	5.892	7.042	9.299	12.340	15.98	19.81	22.36	27.69
14	4.660	6.571	7.790	10.165	13.339	17.12	21.06	23.68	29.14
15	5.229	7.261	8.547	11.037	14.339	18.25	22.31	25.00	30.58
16	5.812	7.962	9.312	11.912	15.338	19.37	23.54	26.30	32.00
17	6.408	8.672	10.085	12.792	16.338	20.49	24.77	27.59	33.41
18	7.015	9.390	10.865	13.675	17.338	21.60	25.99	28.87	34.80
19	7.633	10.117	11.651	14.562	18.338	22.72	27.20	30.14	36.19
20	8.260	10.851	12.443	15.452	19.337	23.83	28.41	31.41	37.57

References

Christensen, Alan C. (2000) Cats as an aid to teaching genetics. *Genetics* 155: 999-1004.

Hardy-Weinberg Problem Set

Note: if you are asked to do a calculation, please show all your work!

1. (4 points) What assumptions are made for Hardy-Weinberg equilibrium to exist?
2. (5 points) Cystic fibrosis is a genetic disorder in homozygous recessive individuals that causes death during the teenage years (although advances in treatment have extended the life expectancy into early adulthood). If 4 in 10,000 newborns have the disease, what are the expected frequencies of the three genotypes in newborns, assuming the population is in Hardy-Weinberg equilibrium? Why is it incorrect to assume Hardy-Weinberg equilibrium in this population?
3. (3 points) The allele p occurs with a frequency of 0.6 in a population of wolves. Again, assuming Hardy-Weinberg give the genotype frequency of PP , Pp , and pp .
4. You are a scientist interested in the population genetics of extinct mammals. Fortunately, someone has finally invented time travel that doesn't require a DeLorean. You decide to take advantage of this technological advance and travel back to the middle of the last ice age in the Midwestern United States to study the woolly mammoth. You find that the color of the fur varies between a reddish brown and a dark brown. After careful observation, you determine that the allele for reddish brown fur is recessive and the allele for dark brown is dominant. You observed 580 dark brown woolly mammoths in a population of 1000. Genetic testing allowed you to determine that 450 of these are heterozygotes.
 - a. (2 points) What is the observed frequency of the homozygous recessive genotype?
 - b. (2 points) Of the animals you observed, how many were homozygous dominant?
 - c. (2 points) What is the frequency of both alleles in the population?
 - d. (4 points) Is this population evolving? Provide quantitative evidence to support your assertion.
 - e. (8 points) You change the setting on time machine allowing you to return a number of centuries later and now find the following: the population size has not changed, but now there are 380 dark brown woolly mammoths. And with genetic testing you determine that 315 are heterozygotes for fur color. Is this population evolving? Provide quantitative evidence to support your assertion.

The Evolution of Caminalcules

Modified from Robert P. Gendron
Indiana University of Pennsylvania

Humans classify almost everything, including each other. This habit can be quite useful. For example, when talking about a car someone might describe it as a 4-door sedan with a fuel injected V-8 engine. A knowledgeable listener who has not seen the car will still have a good idea of what it is like because of certain characteristics it shares with other familiar cars. Humans have been classifying plants and animals for a lot longer than they have been classifying cars, but the principle is much the same. In fact, one of the central problems in biology is the classification of organisms on the basis of shared characteristics. As an example, biologists classify all organisms with a backbone as "vertebrates." In this case the backbone is a characteristic that defines the group. If, in addition to a backbone, an organism has gills and fins it is a fish, a subcategory of the vertebrates. This fish can be further assigned to smaller and smaller categories down to the level of the species. The classification of organisms in this way aids the biologist by bringing order to what would otherwise be a bewildering diversity of species. (There are probably several million species - of which about one million have been named and classified.) The field devoted to the classification of organisms is called **taxonomy** [Gk. taxis, arrange, put in order + nomos, law].

The modern taxonomic system was devised by **Carolus Linnaeus** (1707-1778). It is a **hierarchical** system because organisms are grouped into ever more inclusive categories from species up to kingdom. Figure 1 illustrates how four species are classified using this taxonomic system. (**Important Note:** it is standard practice to underline or *italicize* the genus and specific epithet. The first letter of the genus name is capitalized and the specific epithet is not. The species name includes both the genus name and the specific epithet, ex. *Tyrannosaurus rex*. Please learn to follow this standard practice!)

DOMAIN	Eukarya		
KINGDOM	Animalia		Plantae
PHYLUM	Chordata		Angiospermophyta
CLASS	Mammalia		Monocyledoneae
ORDER	Primate	Carnivora	Hymenoptera
FAMILY	Hominidae	Canidae	Apidae
GENUS	<i>Homo</i>	<i>Canis</i>	<i>Apis</i>
SPECIFIC EPITHET	<i>sapiens</i> (human)	<i>lupus</i> (wolf)	<i>mellifera</i> (honeybee)
			Liliales
			Liliaceae
			<i>Alium</i>
			<i>sativum</i> (garlic)

Figure 1

In the 18th century most scientists believed that the Earth and all the organisms on it had been created suddenly in their present form as recently as 4004 BC. According to this view, Linnaeus' system of classification was simply a useful means of cataloging the diversity of life. Some scientists went further, suggesting that taxonomy provided insight into the Creator's mind ("Natural Theology").

This view of taxonomy changed dramatically when **Charles Darwin** published *On The Origin of Species* in 1859. In his book Darwin presented convincing evidence that life had evolved through the process of **natural selection**. The evidence gathered by Darwin, and thousands of other biologist since then, indicates that all organisms are descended from a **common ancestor**. In the almost unimaginable span of time since the first organisms arose (about 3.8 billion years) life has gradually diversified into the myriad forms we see today.

As a consequence of Darwin's work it is now recognized that taxonomic classifications are actually **reflections of evolutionary history**. For example, Linnaeus put humans and wolves in the class Mammalia within the phylum Chordata because they share certain characteristics (e.g. backbone, hair, homeothermy, mammary glands etc.). We now know that this similarity is not a coincidence; both species inherited these traits from the same common ancestor. In general, the greater the resemblance between two species, the more recently they diverged from a common ancestor. Thus, when we say that the human and wolf are more closely related to each other than either is with a honeybee, we mean that they share a more **recent common ancestor** with each other than either of them does with the honeybee.

Another way of showing the evolutionary relationship between organisms is in the form of a **phylogenetic tree** (Gk. phylon, stock, tribe + genus, birth, origin):

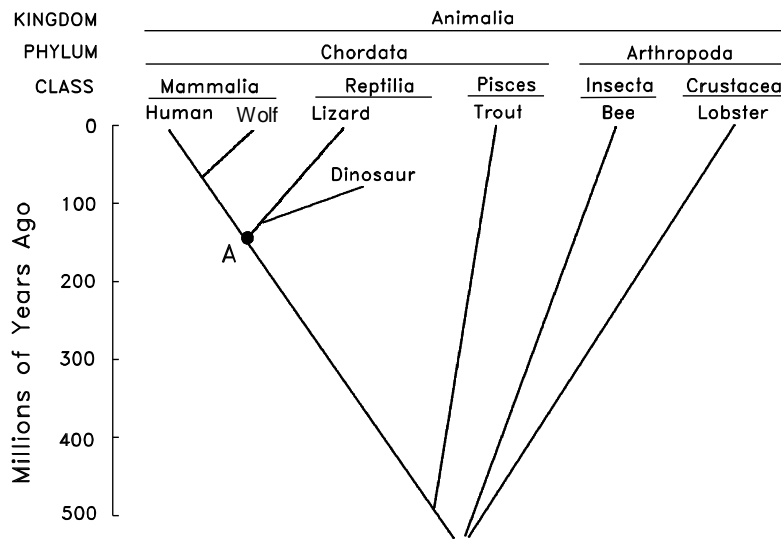


Figure 2

The vertical axis in this figure represents time. The point at which two lines separate indicates when a particular lineage split and is called a **node**. For example, we see that mammals diverged from reptiles about 150 million years ago. The **most recent common ancestor** shared by mammals and reptiles is indicated by the node labeled A. The horizontal axis represents, in a general way, the amount of divergence that has occurred between different groups; the greater the distance, the more different their appearance. Note that because they share a fairly recent ancestor, species within the same taxonomic group (e.g. the class Mammalia) tend to be closer to each other at the top of the tree than they are to members of other groups.

Several types of evidence can elucidate the evolutionary relationship between organisms, whether in the form of a taxonomic classification (Fig. 1) or a phylogenetic tree (Fig. 2). One approach, as already discussed, is to compare living species. The greater the differences between them, the longer ago they presumably diverged. There are, however, pitfalls with this approach. For example, some species resemble each other because they independently evolved similar structures in response to similar environments or ways of life, not because they share a recent common ancestor. This is called **homoplasy**, because distantly related species seem to show the same appearance (become more similar). There are two types of homoplasy, **convergence** and **reversal**. An example of **convergent evolution** would include the wings of bats, birds, and pterosaurs. All of these animals independently evolved wings, and wings are not present in the recent common ancestor. An example of **reversal** would be the streamlined shape of a dolphin and a shark. At first glance it might appear that dolphins are a type of fish. Upon further examination it becomes apparent that this resemblance is superficial, resulting

from the fact that dolphins and fish have adapted to the same environment. The presence of hair, the ability to lactate and homeothermy clearly demonstrate that dolphins are mammals. However, the common ancestor of the dolphin and the shark was an aquatic fish with a streamlined body. The difference is, that at some point fish evolved into terrestrial amphibians, then reptiles, then mammals and finally some mammals returned to the sea as whales and dolphins. The environment selecting for a hydrodynamic shape. Thus they have reversed back to a form that was once present in the common ancestor of sharks and dolphins. However, there are still recognizable differences that tell us of the intervening terrestrial life before the return to the sea, lungs not gills, hair not scales, and milk in the dolphin clearly make it more closely related to terrestrial mammals. Thus, the taxonomist must take into account a whole suite of characteristics, not just a single one.

The fossil record can also be helpful for constructing phylogenetic trees. For example, bears were once thought to be a distinct group within the order Carnivora. Recently discovered fossils, however, show that they actually diverged from the Canidae (wolves, etc.) fairly recently. The use of fossils is not without its problems, however. The most notable of these is that the fossil record is incomplete. This is more of a problem for some organisms than others. For example, organisms with shells or bony skeletons are more likely to be preserved than those without hard body parts.

The Evolution of Artificial Organisms

In this lab you will develop a phylogenetic tree for a group of imaginary organisms called **Caminalcules** after the taxonomist Joseph Camin who devised them. At the back of this lab are pictures of the 14 "living" and 58 "fossil" species that you will use. Take a look at the pictures and note the variety of appendages, shell shape, color pattern, etc. Each species is identified by a number rather than a name. For fossil Caminalcules there is also a number in parentheses indicating the geological age of each specimen in millions of years. Most of the fossil Caminalcules are extinct, but you will notice that a few are still living (e.g. species #24 is found among the living forms but there is also a 2 million year old fossil of #24 in our collection).

The purpose of this lab is to illustrate some of the processes of evolution (e.g. homoplasy). We do these exercises with artificial organisms so that you will approach the task with no preconceived notion as to how they should be classified. This means that you will have to deal with problems such as homoplasy just as a taxonomist would. With real organisms you would probably already have a pretty good idea of how they should be classified and thus miss some of the benefit of the exercise.

The Comparative Approach to Phylogenetic Analysis

When there are three or more species you must decide which two of the species share a common ancestor not shared by the other(s). This diagram indicates that species E and K are more closely related to each other than either is to C. We hypothesize that E and K have a common ancestor (y) that is not shared by C. Similarly, two genera that more closely resemble each other than they do other genera presumably share a common ancestor. We can even infer what a common ancestor like y might have looked like.

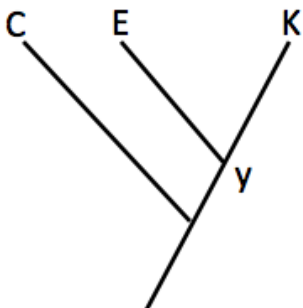
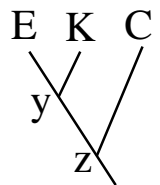


Figure 3

Exercise: The Phylogeny of Caminalcules

Using a large sheet of paper, construct a phylogenetic tree for the Caminalcules. Use a meter stick to draw 20 equally spaced horizontal lines on the paper. Each line will be used to indicate an interval of one million years. Label each line so that the one at the bottom of the paper represents an age of 19 million years and the top line represents the present (0 years).

Cut out all the Caminalcules (including the living species). Put them in piles according to their age (the number in parentheses). Beginning with the oldest fossils, arrange the Caminalcules according to their evolutionary relationship. Figure 4 shows how to get started.

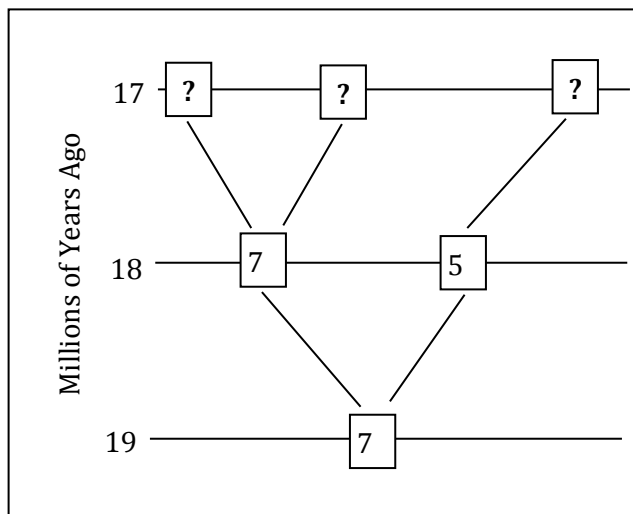
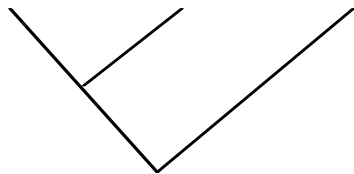


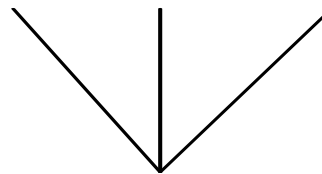
Figure 4: Hypothetical branching pattern of taxa represented by fossils. Based on known taxa and fossils as early as 19 million years ago (mya), #7 is the hypothesized common ancestor of the Caminalcules. Taxa #7 was still present 18 mya, but an additional taxon had evolved as well. What is the hypothesized relationship of the three taxa for which fossils were found 17 mya? Remember that our hypothetical phylogeny is based on the characteristics that we observe in the fossil or living taxa.

Hints, Suggestions and Warnings

- Draw lines faintly in pencil to indicate the path of evolution. Only after your instructor has checked your tree should you tape the figures in place and darken the lines.
- Branching should involve only two lines at a time:



Like this



Not this

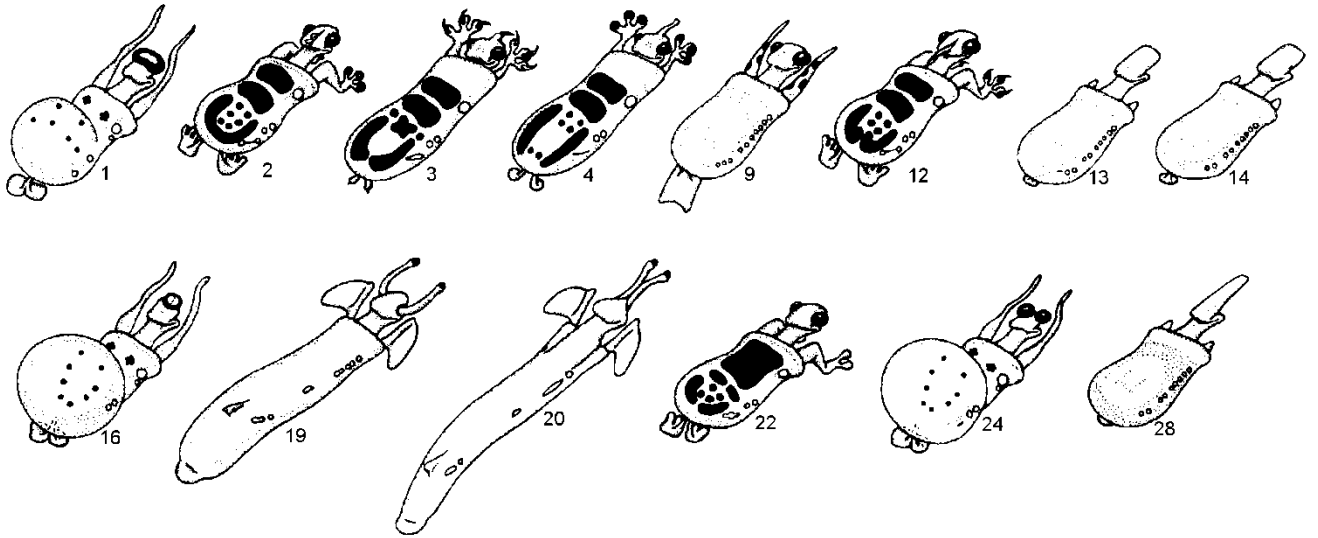
- Some living forms are also found in the fossil record.

- d. There are gaps in the fossil record for some lineages. Also, some species went extinct without leaving any descendants (remember non-avian dinosaurs are extinct).
- e. The Caminalcules were numbered at random; the numbers provide no clues to evolutionary relationships.
- f. There is only one correct phylogenetic tree in this exercise. This is because of the way that Joseph Camin derived his imaginary animals. He started with the most primitive form (#73) and gradually modified it using a process that mimics evolution in real organisms. After you complete your phylogeny compare it with Camin's original tree.

Questions

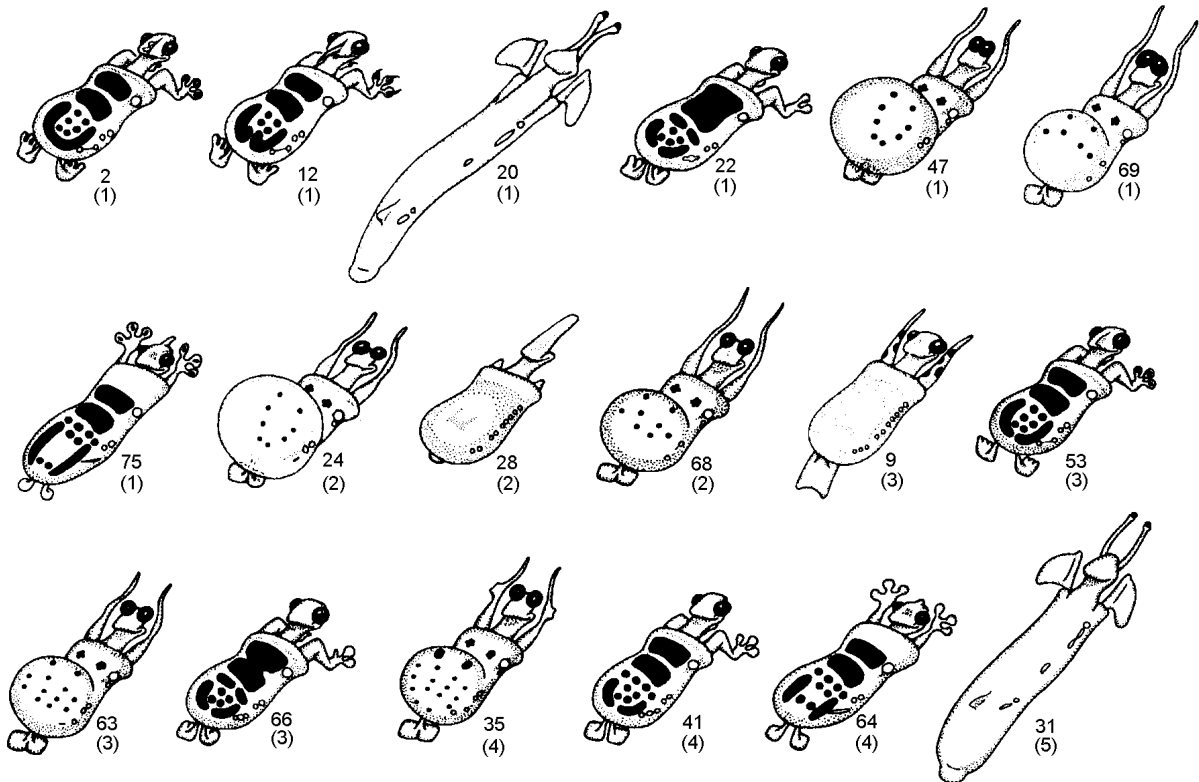
1. How does your phylogeny compare to the actual phylogeny of the caminalcules? Where are your differences. What characters made you chose the pattern you did instead of one like the actual phylogeny.
2. You will notice that some **lineages** (e.g. the descendants of species 56) branched many times and are represented by many living species. Discuss the ecological conditions that you think might result in the rapid diversification of some lineages (A real world example would be the diversification of the mammals at the beginning of the Cenozoic, right after the dinosaurs went extinct.)
3. Some lineages (e.g. the descendants of species 58) changed very little over time. A good example of this would be “living fossils” like the horseshoe crab or cockroach. Again, discuss the ecological conditions that might result in this sort of long-term evolutionary stasis.
4. Some Caminalcules went extinct without leaving descendents. In the real world, what factors might increase or decrease the probability of a species going extinct?
5. Find two examples of **convergent evolution** among the Caminalcules. This means finding cases where two or more species have a similar characteristic that evolved independently in each **lineage**. The wings of bats, birds and pterosaurs is an example of convergence because the three groups did not inherit the characteristic from their **common ancestor**. Write your answers in complete sentences (e.g. “Species x and y both have ____ but their most recent common ancestor, z, did not”).
6. Describe two examples of **vestigial structures** that you can find among the Caminalcules. These are structures that have been reduced to the point that they are virtually useless. Ear muscles and the tail bones are examples of vestigial structures in our own species. Explain how vestigial structures provide clues about a species’ evolutionary past. Illustrate your argument with vestigial structures found in humans or other real species

LIVING CAMINALCULES

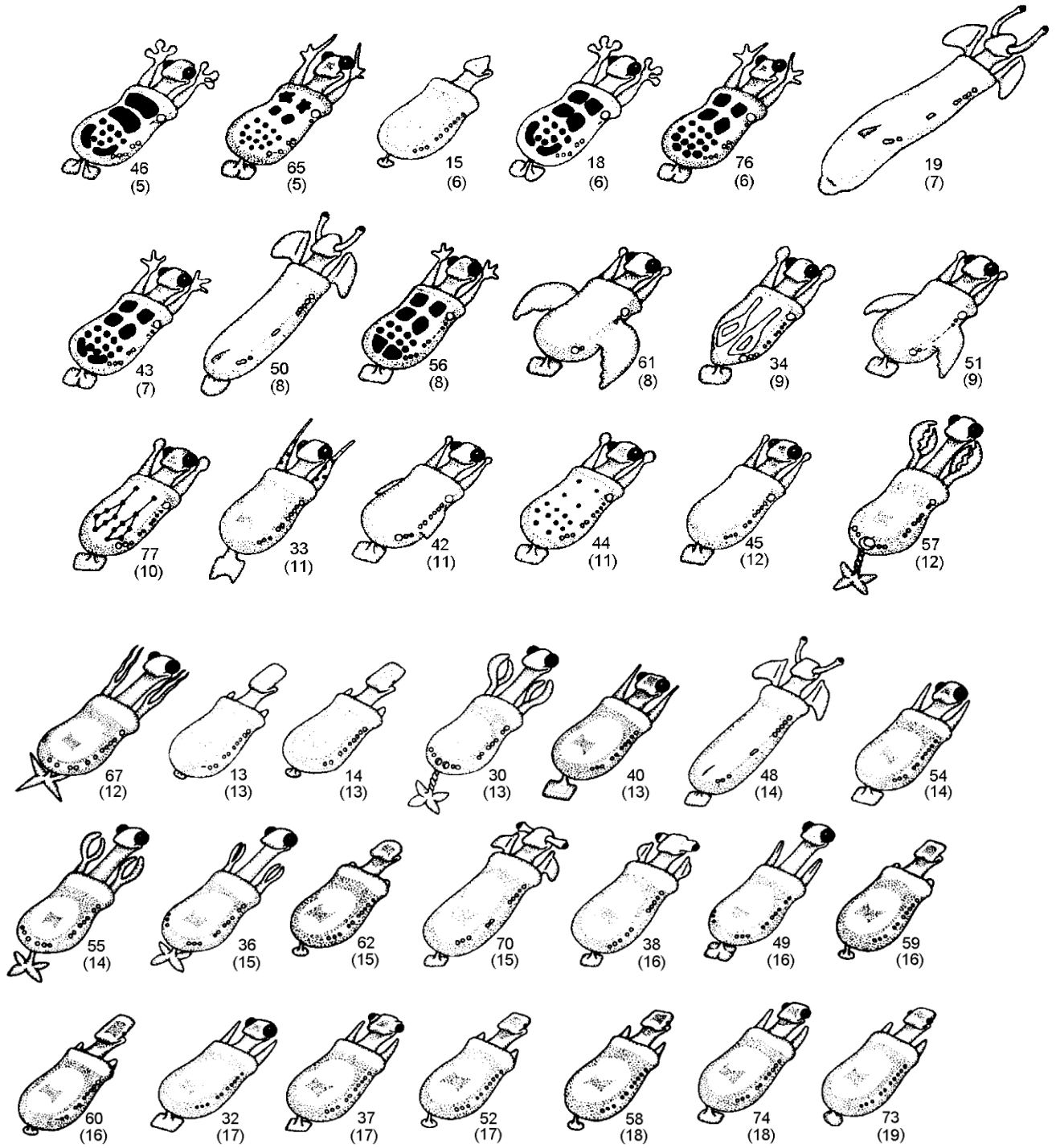


FOSSIL CAMINALCULES

(numbers in parentheses indicate age in millions of years)



FOSSILS (continued)

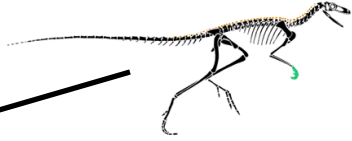


Review of Phylogenetic Trees

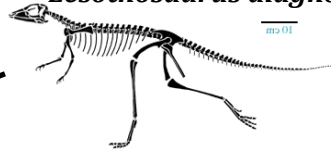
- 1) Using the dinosaur phylogeny and the character-taxon matrix on the following pages, map the characters onto the phylogeny. How many steps long is this phylogeny? Note: *Lagosuchus talanpayenisis* is the outgroup for this phylogeny.
- 2) Name all characters that are homoplastic based on your phylogeny and your reasoning for deeming them so.
- 3) What character or characters are shared derived for the recent common ancestor of modern birds, *Lesothosaurus diagnosticus* and all of its descendants?
- 4) Feathers, hollow bones, and a semi-lunate carpal are all characters that allow modern birds to function in both airborne and terrestrial locomotion. Hollow bones give birds a light but strong skeleton, a semi-lunate carpal allows bird to tuck their wings next to their body when they are not flying, and most importantly, feathers act as an airfoil allowing for aerodynamic lift. Based on your phylogeny, did birds evolve these characters because they needed them to fly? Justify your answer using your phylogeny and be thorough in your argument.
- 5) Are there any other character/characters not mentioned in question 4 but given in the matrix that would have facilitated the evolution of flight, allowing birds to gain this ability. Why do you think so?
- 6) Exaptations are adaptations that were shaped by natural selection for a particular function that is different than their current function. In other words, the character in question served a different function than the one it is currently used for. What function do you think feathers served dinosaurs prior to their use as airfoils in modern birds? Hint: What else do modern birds use their feathers for?
- 7) Birds have been considered by some to be more closely related to the clade that consists of *Lesothosaurus*, *Scutellosaurus*, and *Stegosaurus*. Redraw the phylogeny so that modern birds are the sister group to these taxa and plot the characters on this new phylogeny. Is this hypothesis of relationships better supported than the first? Why or why not?

Character/taxon matrix	Open Acetabulum	Quadrupedal or Bipedal	Heavy Armor	Opisthopubic Pelvis	Hollow bones	Feathers	Semi-lunate Carpal Bone	Flight
(OUTGROUP) <i>Lagosuchus talanpayenisis</i>	Absent	Bipedal	Absent	Absent	Absent	Absent	Absent	Absent
<i>Lesothosaurus diagnosticus</i>	Present	Bipedal	Absent	Present	Absent	Absent	Absent	Absent
<i>Scutellosaurus lawleri</i>	Present	Bipedal	Present	Present	Absent	Absent	Absent	Absent
<i>Stegosaurus stenops</i>	Present	Quadrupedal	Present	Present	Absent	Absent	Absent	Absent
<i>Plateosaurus englehardti</i>	Present	Bipedal	Absent	Absent	Present	Absent	Absent	Absent
<i>Diplodocus carnegii</i>	Present	Quadrupedal	Absent	Absent	Present	Absent	Absent	Absent
<i>Ceratosaurus nasiconus</i>	Present	Bipedal	Absent	Absent	Present	Absent	Absent	Absent
<i>Caudipteryx zoui</i>	Present	Bipedal	Absent	Absent	Present	Present	Absent	Absent
<i>Sinornithosaurus millennii</i>	Present	Bipedal	Absent	Present	Present	Present	Present	Absent
<i>Archaeopteryx lithographica</i>	Present	Bipedal	Absent	Present	Present	Present	Present	Present
Modern Birds	Present	Bipedal	Absent	Present	Present	Present	Present	Present

Lagosuchus talanpayenensis



Lesothosaurus diagnosticus



Scutellosaurus lawleri



Stegosaurus stenops



Plateosaurus engelhardti



Diplodocus carnegii



Ceratosaurus nasiconus



Caudipteryx zoui



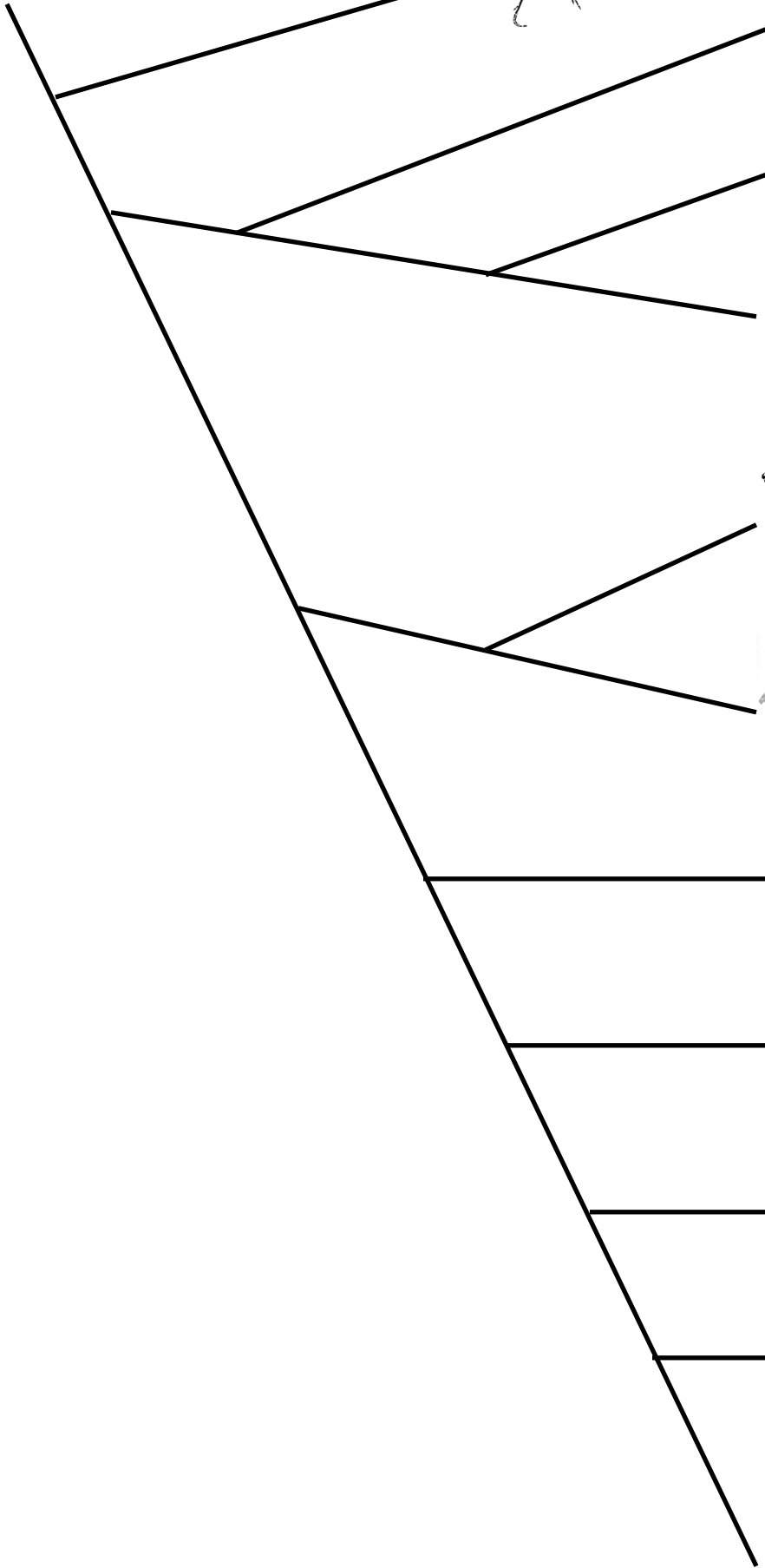
Sinornithosaurus millennii



Archaeopteryx lithographica



Modern Birds



Phylogenetic Systematics Lab: Evolution of Ceratopsian Dinosaurs

Biol 113: Evolution & Diversity

Frank J. Varriale

Materials

- Images of dinosaur skulls, statistics programs with phylogenetic software and analysis tools (PAST and Mesquite). If the appropriate software is not installed on your computer you can download and install these shareware programs from the below sites.
- **PAST**(PALaeontologicalSTatistics)
<http://folk.uio.no/ohammer/past/download.html>
Click on the **Past.exe** file and install from there.
- **Mesquite**
<http://mesquiteproject.org/mesquite/download/download.html>
Download and install the **Windows version**
- **Java** latest version (only if you can't see all the icon buttons in Mesquite)
<http://www.java.com/en/download/>

Introduction

Ceratopsians are popularly referred to as the horned-dinosaurs because many members of this group have large horns protruding from above their eyes, nose, and occiput. Indeed, the word “Ceratopsia” means horned-face in Greek. Some ceratopsians have a large frill that looks like a shield extending from the back of the head. The oldest known ceratopsian is *Yinlong downsi* from the Late Jurassic period of China, approximately 162 million years ago (mya). The last known ceratopsian was the iconic *Triceratops horridus*, which lived in the western United States during the Late Cretaceous period until 65mya. It is one of the last non-avian dinosaurs, and became extinct due to the Cretaceous-Tertiary extinction event that killed all dinosaurs except birds. Ceratopsians had an impressive 97 million year existence. Considering that most of the large mammal diversity has occurred during only the last 65 million years, ceratopsians had a great length of time in which to experience evolutionary change.

Goals

- Acquire a basic mastery of the terminology and methodology of phylogenetic systematics.
- Understand how species are related to each other and how characters are used to determine those relationships.

Instructions

Part 1 (Coding dinosaur characters)

- 1) Look at the provided images of ceratopsian dinosaurs and the outgroup (*Hypsilophodon foxii*) to get an appreciation for the morphologic diversity in skull shape between these dinosaurs.
- 2) Each of you will be assigned at least one or two characters that you will score for all, or a subset of the ceratopsians, your instructor will let you know. You will use the provided character state list to score the dinosaurs.

Important note: There are 14 images of ceratopsian dinosaurs and one outgroup taxon (*Hypsilophodon foxii*). As you will remember from class and your reading, the **outgroup** is the taxon or taxa that you are comparing your group of interest to. The group you are interested in, and whose evolutionary relationships you want to understand is called the **ingroup**. In this lab the ingroup is the ceratopsian dinosaurs. The outgroup is assumed to possess the **ancestral** or **plesiomorphic** character states because it branched off of the tree at a different point than the members of your ingroup. You use the outgroup to help you understand which states are the ancestral condition so that you can see how these characters change in your group of ceratopsians. This allows you to understand **polarity**, the basic direction of change. Any

character that is shared between two taxa and their common ancestor is called a **synapomorphy**, and also known as a shared derived character. It is the sharing of these characters that tells you which organisms are more closely related to each other. To be confident that the ancestral states are truly ancestral (and not unique characters that evolved independently in your outgroup = **autapomorphies**) you would include a second outgroup to provide more certainty about the ancestral condition. However, for the sake of brevity you will only be scoring one outgroup which is known to possess the ancestral states.

- 3) Take out a sheet of paper and make a **character/taxon matrix** for the character(s) and taxa that you have been assigned. Taxa are listed in the left most column. Characters are listed in the top row and the matrix is filled in based on the provided description of the different character states and their associated numbers. Remember that your outgroup taxon goes at the top of the list and that the character states the outgroup possesses receive a score of 0 (zero); because this is the ancestral state, and anything different from the outgroup is considered to be **derived** relative to the ancestral condition. Derived character states receive a coding of 1, or in the case of characters that have multiple different states, each character state gets its own number (1,2,3 *etc.*).
- 4) When you are finished scoring your characters, go to the computer at the podium and type in your recorded states for each of the characters you scored. The character taxon matrix at the podium has all the character #'s listed in the first row. Make absolutely certain that you put your recorded characters in the correct numbered column and that your taxa are in the same order!
- 5) After the entire class has done this, I will post the file and you will download it and analyze it using the PAST and Mesquite software next time we meet.

Part 2 (Analyzing Trees)

- 6) If it is not already installed, download and install the PAST and Mesquite software from the sites listed above. On the PAST website click and run the past.exe file. On the Mesquite website chose the “Windows” option.
- 7) We will look at a simple example first before we analyze your ceratopsian matrix. This will give you an idea of how to use the program and how to interpret the display and summarized statistics.
- 8) The example we will use is the same matrix (Table 2.1, p.16) that is in your reading from the Berta & Sumich (1999) systematics and classification chapter (2) of marine mammals.
- 9) In the spreadsheet page of PAST make sure that both the “edit mode” and “edit labels” buttons are checked so that you can label the rows and columns. Label them based on the same labels in Table 2.1, but abbreviate them so that they don’t run out of the space provided and there is only one word as a simple description. Use the word bears as a substitute for “outgroup” because bears are the outgroup. **Do Not put blank spaces in your labels or this will create problems later.**
- 10) In order for the computer to analyze these data, the verbal coding in Table 2.1 needs to be converted to numerical coding. So..., how are you going to do that? What number do we give the character states of the outgroup? Notice that character two (Locomotor type) is a multistate character and each state receives its own numerical code. What are your thoughts on how we should code this character?
- 11) When you are done creating the character/taxon matrix, go to Edit→Select All and a blue highlight should appear over all your data.
- 12) Go to Cladistics→Parsimony analysis and a dialogue window should appear. Make sure that the “branch-and-bound” algorithm and the “Fitch” optimization are checked, and leave the rest of the options at their default values.
- 13) Then hit Go! **Important:** do not do anything with the computer while it is analyzing data or the program will crash. So... sit back and wait until it is finished computing most parsimonious trees. This should go quickly but when you do the ceratopsians it will take longer.
- 14) In the original Parsimony analysis window it indicates that the program analyzed three possible trees (Trees evaluated), these correspond to the three trees; b, c, and d in Fig 2.4 of the reading. It also shows that there was only one most parsimonious tree (**# trees stored: 1**) and that this tree had six steps (**Shortest tree: 6**)

- 15) In the new “Most parsimonious trees” window the first of all the most parsimonious trees is shown. This tree can also be referred to as a **cladogram** because it shows the relationships of various groups (**clades**) of organisms due to the sharing of derived characters (**synapomorphies**). In this example there is only one tree, but if there were more you could move through them by clicking the up/down arrows next to “Tree 1”. This window also shows the tree length as six steps. Wherever there is a branching event that is called a **node**, at that node is the theoretical **recent common ancestor** of the two taxa on the branches extending away from that node. For instance, in your tree, Walruses and Seals share a more recent common ancestor with each other than either does with Fur Seals + Sea lions or bears. Because of this, Walruses and Seals are said to be more closely related to each other.
- 16) Go down to “branch lengths” and click that button. Now the lengths of the branches appear in blue showing you the number of evolutionary steps that are located on that branch. In total there are six, as you might expect from the total tree length. However, we can see that two of these steps occurred on the branch leading to Fur Seals + Walruses + Seals while the other four steps are only on the branch leading to Walruses + Seals. Look at Figure 2.4b and verify that your analysis and this figure both have six steps and that those steps are located in the same place on the tree.
- 17) You can also see how each of the characters is distributed on the tree by clicking the up arrow next to “Character NONE”. Character 1; Thick fat present, shows up once on the tree in the branch leading to Walruses + Seals and is found in both of these taxa. Run through all the characters and examine their distributions on the tree.
- 18) Go back to the NONE Character and then click the middle box of the three boxes that are under the “Consensus” box. You have now copied the figure. Open Microsoft Word and paste it into a new document and save it so that you can have it to print later for the lab turn in.
- 19) Now go back to PAST and go to File → Save as... and in the “Save as type”, save the file as a NEXUS (.nex) file and title the file “Pinnipeds”. This will allow you to open the file in the Mesquite program.
- 20) Open the Mesquite program and then the file “Pinnipeds” by using the file → open file in the top left. Mesquite will take a few seconds, so wait for it. A window with lots of options will appear. Go to the bottom left and click on the icon that says “View Trees” next to it. This will open an area for you to view and analyze your trees.
- 21) In the menu at the top of the new window with the tree, go to Drawing → Tree Form → Diagonal tree and choose it. Then go to Analysis → Trace character history, and choose the default value “Stored Characters”, then click OK, and choose “Parsimony Ancestral States” and finally click OK again.
- 22) A small legend appears in the lower corner and you can click the arrow to move through the characters and see the evolutionary history of their appearance on the tree. Multistate characters are colored coded. Look through all five and appreciate the evolutionary history of their appearance.
- 23) Scroll back to Character 2 (Locomotor type). Now look at the branch leading to Fur seals and Walrus + Seals it should be both black and white; this is because there are multiple equally parsimonious options for the character states on this branch. The “front limbs used” in swimming could have appeared here and then evolved into the “hind limbs used” that is characteristic of Walrus + Seals, or it could have appeared higher up on the tree in the branch leading to Fur seals and Sea lions. The first possibility is called **accelerated transformation**, because the change comes on the tree earlier in time, thus it is accelerated. The second possibility is called **delayed transformation** because the change occurred later on the tree only in the branch leading to Fur Seals. Now put your pointer over the top of this branch and notice that in the “Trace character” area it says most parsimonious states 0, 1, 2. This is because that portion of the branch could have the states 0, 1, or 2 and with your current data there is no way to know which is the real evolutionary history of this character, because each of these three options still yields a tree with 6 steps, thus they are all equally parsimonious, and equally short.

Important note: Even here we are employing the scientific method. Each possible tree and the character states plotted on them are alternative hypotheses about the evolutionary relationships and history of a group of organisms. You test each hypothesis by using the concept of parsimony. **Parsimony** is the preference for the least complex explanation and thus the least amount of assumptions. In the case of

phylogenetic systematics, the trees with the least amount of assumptions are the ones with the least amount of evolutionary changes or steps. These trees are the best supported alternate hypotheses about how an evolutionary history unfolded. However, when you have multiple equally parsimonious trees, your data are unable to help you reject one tree (hypothesis) and accept another.

- 24) Now go to Analysis → Values for Current Trees... and in the window that appears chose the default “Tree value using character matrix” and click OK. Then in the next window chose “Treelength” and click ok, and then in the final window keep the default “Stored Matrices” and again click OK. A small blue legend will appear. Confirm that it indicates your treelength is six steps.
- 25) Now use your arrow and click on and hold the Fur seal/Sea lion (FS/SL) branch and drag it to the branch going to seals. Once you have extended the line to the “Seals” branch, let go and you will have transformed the tree to one where FS/SL are now more closely related to Seals.

Question 1: What happened to the tree length and what does this mean for the new tree compared to the original.

Question 2: Now scroll through the characters to find out why your answer in question 1 is this way. Look at Character 2, now there is no ambiguity in where the states of this character appear on the tree. However, what happened to Characters 1, 3, and 4?

Question 3: How many times does the derived state for each character 1, 3, and 4, appear on the new tree compared to the original tree. What do we call this pattern of appearance? Hint: look in the reading on marine mammals I gave you.

Part 3 (Ceratopsian Evolution)

- 26) Go to Moodle and get the Ceratopsians nexus file we created from last time.
- 27) Open it up in PAST and run the Cladistic analysis like you did with the marine mammals. This may take several minutes, **do not** work on the computer or you will crash the program. Just wait till it is finished comparing trees.
- 28) When the analysis is finished take note that there will be more than 1 most parsimonious tree (MPT). As a side note, appreciate how many trees the computer evaluated while you were waiting? Click the Consensus option at the bottom of the window and chose the default strict consensus. A **strict consensus tree** is created by constructing a tree from only those clades (relationships) that were present in every one of the MPTs. Click Majority and see what changes. A **majority rule tree** is created from those clades that are the same across a particular percentage of all the MPTs. In some programs you may be able to specify this percent however in PAST this option is preprogrammed for only those clades that appear in 50% or greater of all the MPTs.
- 29) Go back to the strict consensus tree and click the middle of the three boxes that are under the Consensus options and paste this file into the MS Word document you created earlier and save it. Before you close anything, save your file again as a nexus file!, this will preserve your trees that you found, along with the other data so that you can analyze it with Mesquite. Close PAST
- 30) In Mesquite open the Ceratopsians file like you did earlier with the marine mammals in Part 2 #20.
- 31) Once in the view trees you need to get the consensus tree. Go to Tree → Tree Source → Consensus. In the window click the default “Stored Trees”, then OK, and then “Strict Consensus” and OK again. In the next window open the pull-down menu and choose “Rooted”.
- 32) Make sure you do the same procedure in Part 2 steps 21 and 24 to get back the legends and settings you were once using in the marine mammal example.

Question 4: Compare your ceratopsians tree in Mesquite to the one that you saved in the MS Word document. Are they the same? Why or why not? Does this mean that the two programs (PAST & Mesquite) are showing you different relationships for the various ceratopsians? Explain!

Question 5: Compare your tree to the complete tree of all known ceratopsians provided in this lab. This tree is a super-tree created by combining several analyses that used many more characters and taxa than you did. Did you find in your analysis the same major clades that are named at nodes in the complete phylogeny? For instance, do you have a group that you could call Centrosaurinae. Name all the major groups that are the same.

Question 6: Do you have any **polytomies** (multifurcations) in your tree? If you do have them what do you think caused them to appear? Do they really indicate multiple origins of several taxa at the same time, Justify this answer?

Question 7: List all the character states that are synapomorphies for the major ceratopsian clades that are in your analysis. For example, Ceratopsidae, Centrosaurinae, Chasmosaurinae. What character state or states are common (synapomorphic) to all ceratopsians? You do not have to write the name of the character and the state just write the character number and state number like this 1(2) would mean character 1 state 2 is a synapomorphy.

Question 8: Characters that are unique to a single taxon are called **autapomorphies**. Are there any in your analysis? If you found some, name the taxa and the character state that are autapomorphic. Are these characters helpful in telling you anything about the relationships between groups. Why?

The marine mammal chapter 2 mentions two types of **homoplasy**, one type is due to **convergence** and the other is the result of **reversal**. Remember, homoplasy is when two taxa or different groups share a character but that character is not present in their common ancestor.

Question 9: List all the character states that appear to be due to reversal. How were you able to identify these characters as reversals?

Question 10: List all the character states that show clear evidence of convergence. How were you able to identify these characters as convergent?

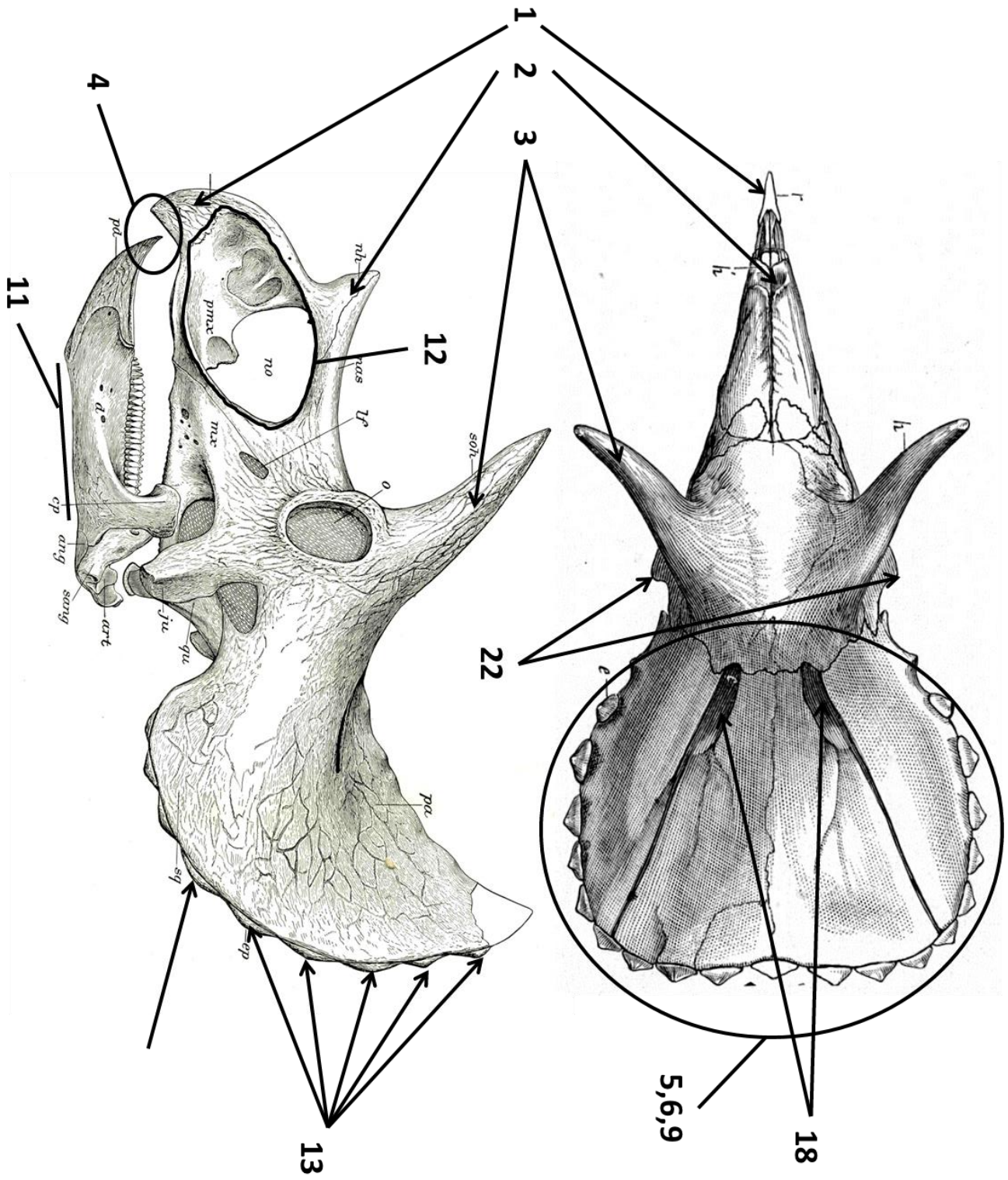
33) Now we are going to address some paleobiogeographical issues. In Mesquite, look to the left and click the “show matrix” icon. In the tools menu to the left of the matrix click the add characters tool. Put your mouse over the icons and a description will appear near the bottom of the window. Use it to add a character 27 at the end of 26, by clicking it and then clicking in the space at the end of the matrix. When it says “How many characters to add” just do one. This will be a character for geographic locality and the states of this character will be the continents that your dinosaurs come from. Use Eurasia as one state, and North America as another. How are you going to code this? Remember the outgroup gets a zero (0) for its state because it is ancestral. Use the provided map to fill in the states for your taxa. Click the “edit” icon and then you can put your mouse in the character state box and type in it.

34) Once you are finished, go back to your consensus tree by clicking the tab at the top of the window called “Tree Window 1 showing Consensus Tree” and then scroll to the new 27th character to see how the new geographic locations plots on the tree. Then answer the following question.

Question 11: What continent did Ceratopsia originate on? Does the cladogram indicate that ceratopsians migrated to another continent? What continent did they migrate to? Justify this answer using the phylogeny.

Print and turn in all your copied MS Word phylogenies and the answers to your questions when your professor asks for them!

Ceratopsian Character Guide



Ceratopsian Character list

- 1) Rostral bone (bone at tip of snout)
 - 0 Absent
 - 1 Present

- 2) Nasal horn (horn at the tip of nose)
 - 0 Absent
 - 1 Small bump
 - 2 Short horn
 - 3 Long horn that is straight or curved
 - 4 No horn but there is lots of rough ridges on its surfaces

- 3) Horns over eyes
 - 0 Absent
 - 1 Long horns
 - 2 Short horns
 - 3 Rough bumps over the eyes

- 4) Shape of the beak
 - 0 Flat/slightly pointy
 - 1 Pointy at the tip like a hawk or turtle

- 5) Frill (sticks out the back of head)
 - 0 Absent = can you see the neck joint (occipital condyle) in dorsal view? If yes then there is no frill.
 - 1 Present

- 6) Frill length
 - 0 Absent
 - 1 Less than or equal to the distance from the eye to the tip of the snout
 - 2 Greater than the distance from the eye to the snout tip

- 7) Parietal fenestra (holes in parietal bone of frill as viewed from above)
 - 0 Absent
 - 1 Present

- 8) Shape of parietal fenestra (holes in frill as viewed from above)
 - 0 Absent
 - 1 Present but circular/somewhat oval
 - 2 Present but elongated ovals in the mediolateral (right to left) direction
 - 3 Present but very elongated ovals in the rostrocaudal (front to back) direction

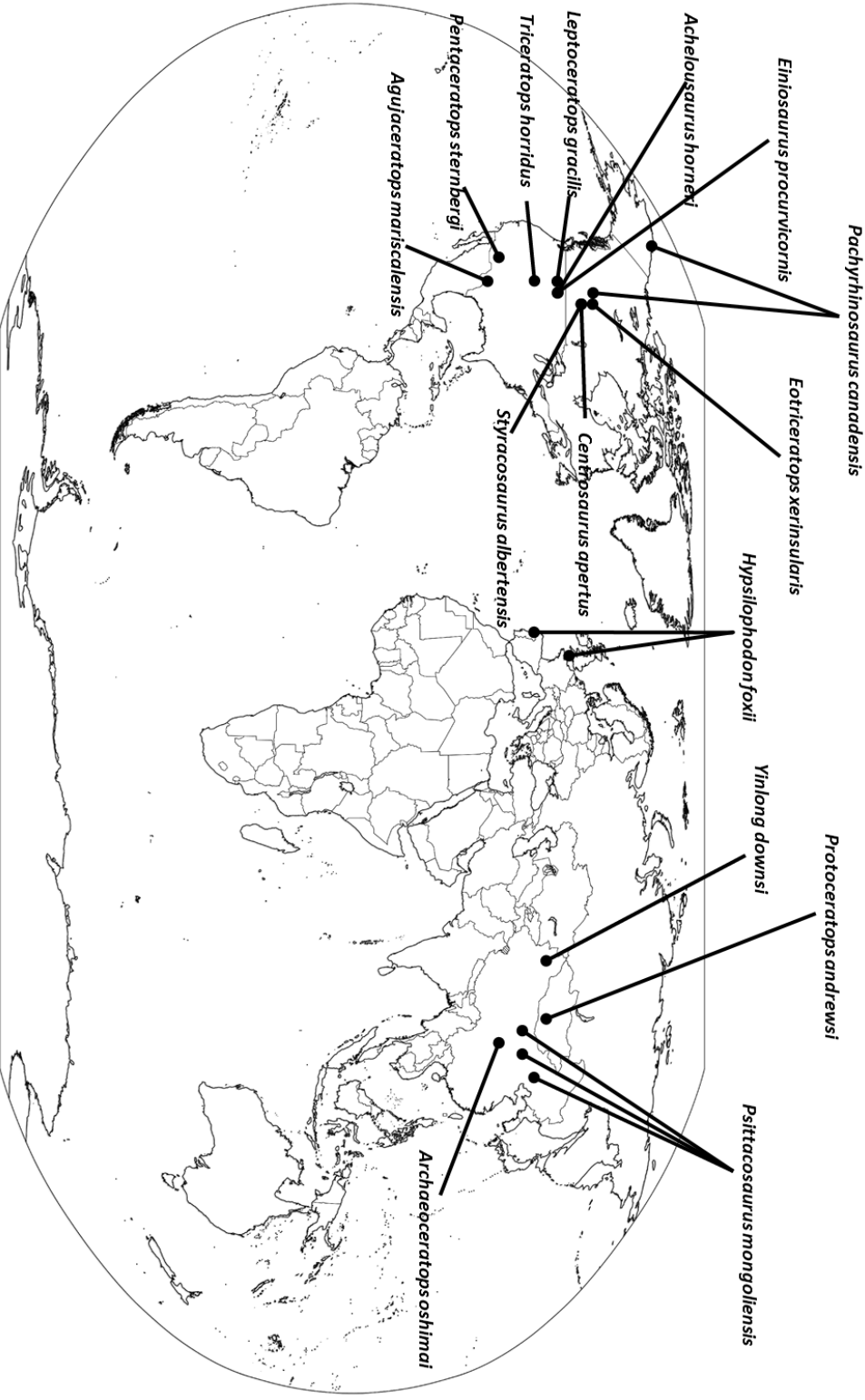
- 9) Shape of the frill in dorsal view without epoccipital spikes (spikes are the little scallops or horns at edge of frill)
 - 0 Absent or not enough frill to show a shape
 - 1 Triangular outline
 - 2 Semicircular outline
 - 3 Rectangular or box shaped outline

- 10) Quadratojugal bone = it is the bone at the tip of the cheek just above the jaw joint
 - 0 Visible in side view
 - 1 Covered by the jugal bone in side view

- 11) Ventral border of the denary and predatory bones in the lower jaw (**Note: Use *Triceratops* for taxa 12-14**)
- 0 Straight
 - 1 Slightly convex ventrally (curved downward)
 - 2 Strongly convex ventrally
- 12) External Naris (Hole for nose)
- 0 Smaller than eye hole (orbit)
 - 1 Larger than eye hole (orbit)
- 13) Epoccipitals (small bones and horns that rim the edge of the frill and give it scalloped edges)
- 0 Absent
 - 1 Present but all nearly the same size
 - 2 Present but some are slightly larger than others
 - 3 Present but some are much larger than others and look like spikes.
- 14) Shape of the second pair of epiparietal spikes (Large spikes at the back edge of the frill)
- 0 Absent
 - 1 Curved forward
 - 2 Long and can be curved to the side and sticking out the back of the frill
- 15) First pair of epiparietal spines
- 0 Absent
 - 1 Directed toward the midline
 - 2 Directed forward
 - 3 Present but not directed as in 1 and 2
- 16) Size of first pair of epiparietal spines
- 0 Absent
 - 1 Same size as the second pair
 - 2 Smaller than the second pair
 - 3 Larger than the second pair
- 17) Small tab of nasal bone sticking into the posterior margin of the external naris (nose hole)
- 0 Absent
 - 1 Present
- 18) Supratemporal fenestra = holes on the top of the skull just behind the eyes. (need to look at top of skull)
- 0 Circular
 - 1 Rectangular to Triangular
 - 2 slit shaped
- 19) Premaxillary teeth
- 0 Present
 - 1 Absent
- 20) Lacrimal foramen (small hole in front of the eye)
- 0 Present
 - 1 Absent
- 21) Lacrimal foramen shape (small hole in front of the eye)
- 0 Larger than the nose hole
 - 1 Smaller than the nose hole
 - 2 Absent but there is a depression there
 - 3 Absent and no depression

- 22) Cheek bones (jugal) shape as viewed from above
- 0 Rounded and do not have a distinct point
 - 1 Come to a distinct point that pokes out laterally
- 23) Posterior margin of frill
- 0 Absent
 - 1 Straight
 - 2 Shallow indentation at the midline
 - 3 Large notch indentation
 - 4 Small notch indentation
- 24) Tab of bone in the premaxilla at the anterior and ventral edge of the nose hole (external naris)
- 0 Absent
 - 1 Present
- 25) Epoccipital bones on the margin of the squamosal bone
- 0 Absent
 - 1 Present
- 26) Notch in the ventral border of the squamosal bone
- 0 Absent
 - 1 Present

Global Map of Ceratopsian Discoveries



Marine Mammals: Evolutionary Biology

Annalisa Berta

*Department of Biology
San Diego State University
San Diego, California*

James L. Sumich

*Biology Department
Grossmont College
El Cajon, California*

With illustrations by

Pieter Arend Folkens

*A Higher Porpoise Design Group
Benecia, California*

Peter J. Adam

*Department of Biology
San Diego State University
San Diego, California*



ACADEMIC PRESS

A Harcourt Science and Technology Company

San Diego San Francisco New York Boston London Sydney Tokyo

2

Systematics and Classification

2.1. Introduction: Systematics—What Is It and Why Do It?

Systematics is the study of biological diversity that has as its emphasis the reconstruction of **phylogeny**, the evolutionary history of a particular group of organisms (e.g. species). It also includes taxonomy, the description, identification, and classification of species. Although the taxonomy of mammals is relatively well known compared to other groups of organisms, we still are discovering previously unknown species (e.g. two new species of beaked whale were discovered in the early 1990s; Reyes *et al.*, 1991, 1995).

Systematic knowledge provides a framework for interpreting biological diversity. Because it does this in an evolutionary context it is possible to examine the ways in which attributes of organisms change over time, the direction in which attributes change, the relative frequency with which they change, and whether change in one attribute is correlated with change in another. It also is possible to compare the descendants of a single ancestor to look for patterns of origin and extinction or relative size and diversity of these groups. Systematics also can be used to test hypotheses of adaptation. For example, consider the evolution of the ability to hear high frequency sounds, or **echolocation**, in toothed whales. One hypothesis for how toothed whales developed echolocation suggests that the lower jaw evolved as a unique pathway for the transmission of high frequency sounds under water. However, based on a study of the hearing apparatus of archaic whales, Thewissen *et al.* (1996) proposed that the lower jaw of toothed whales may have arisen for a different function, one of transmitting low frequency sounds from the ground, as do several vertebrates including the mole rat. According to this hypothesis, the lower jaw became specialized later for hearing high frequency sound. In this way the lower jaw of toothed whales may be an **exaptation** for hearing high frequency sounds. An exaptation is defined as any adaptation that performs a function different from the function that it originally held. A more complete understanding of the evolution of echolocation requires examination of other characters involved such as

the presence of a melon and the morphology of the middle ear and jaw as well as the bony connections between the ear and skull (see Chapter 10).

An understanding of the evolutionary relationships among species can also assist in identifying priorities for conservation (Brooks *et al.*, 1992). For example, the argument for the conservation priority of sperm whales is strengthened by knowing that this lineage, which consists of two extant species, occupies a key phylogenetic position as the basal relative to the other species of toothed whales. These pivotal species are of particular importance in providing baseline comparative data for understanding the evolutionary history of the other 65 species of toothed whales. Sperm whales provide information on the origin of various morphological characters such as suction feeding and the adaptive role of these features in the early evolution of toothed whales.

Perhaps most importantly, systematics predicts properties of organisms. For example, as discussed by Promislow (1996), it has been noted that some toothed whales (e.g. pilot whales and killer whales) which have extended parental care also show signs of reproductive aging (i.e. pregnancy rates decline with increasing age of females), whereas baleen whales (e.g. fin whales) demonstrate neither extended parental care nor reproductive aging (Marsh and Kasuya, 1986). Systematics predicts that these patterns would hold more generally among other whales and that we should expect other toothed whales to show reproductive aging.

Finally, systematics also provides a useful foundation from which to study other biological patterns and processes. Examples of such studies include the coevolution of pinniped parasites and their hosts (Hoberg, 1992, 1995), evolution of locomotion and feeding in pinnipeds (Berta and Deméré, 1993; Adam and Berta, 1997), evolution of body size in phocids (Wyss, 1994), evolution of phocid breeding patterns (Perry *et al.*, 1995), and the evolution of echolocation in toothed whales (Thewissen *et al.*, 1996).

2.2. Some Basic Terminology and Concepts

The discovery and description of species and the recognition of patterns of relationship among them is founded on the concept of evolution. Patterns of relationships among species are based on changes in the features or **characters** of an organism. Characters are diverse, heritable attributes of organisms that include DNA base pairs, anatomical and physiological features, and behavioral traits. Two or more forms of a given character are termed the **character states**. For example, the character “locomotor pattern” might consist of the states “alternate paddling of the four limbs (quadrupedal paddling),” “paddling by the hind limbs only (pelvic paddling),” “lateral undulations of the vertebral column and hind limb (caudal undulation),” and “vertical movements of the tail (caudal oscillation).” Evolution of a character may be recognized as a change from a preexisting, or **ancestral** (also referred to as **plesiomorphic** or **primitive**), character state to a new **derived** (also referred to as **apomorphic**) character state. For example, in the evolution of locomotor patterns in cetaceans, the pattern hypothesized for the earliest whales is one in which they swam by paddling with the hind limbs. Later diverging whales modified this feature and show two derived conditions: 1) lateral undulations of the vertebral column and hind limbs and 2) vertical movements of the tail.

The basic tenet of **phylogenetic systematics**, or **cladistics** (from the Greek word meaning “branch”), is that shared derived character states constitute evidence that the species possessing these features share a common ancestry. In other words, the shared derived fea-

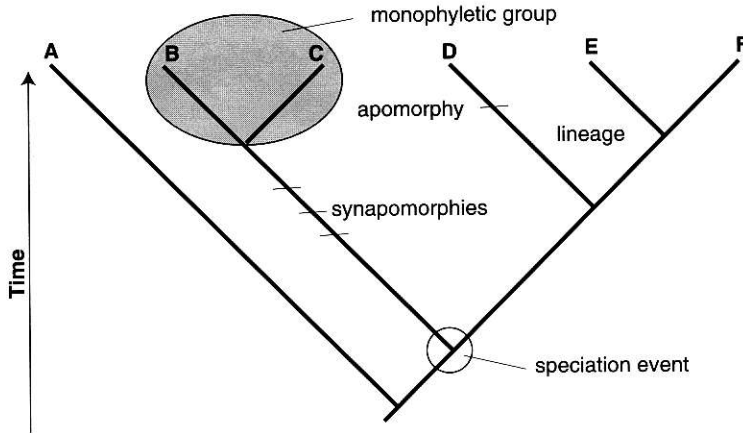


Figure 2.1. A cladogram illustrating general terms discussed in the text.

tures or **synapomorphies** represent unique evolutionary events that may be used to link two or more species together in a common evolutionary history. Thus, by sequentially linking species together based on their common possession of synapomorphies, the evolutionary history of those taxa can be inferred.

Relationships among taxonomic groups (e.g. species) are commonly represented in the form of a **cladogram**, or **phylogenetic tree**, a branching diagram that conceptually represents the best estimate of phylogeny (Fig. 2.1). The lines of the cladogram are known as **lineages** or **clades**. Lineages represent the sequence of ancestor–descendant populations through time. Branching of the lineages at **nodes** on the cladogram represents **speciation events**, a splitting of a lineage resulting in the formation of two species from one common ancestor.

The task in inferring a phylogeny for a group of organisms is to determine which characters are derived and which are ancestral. If the ancestral condition of a character or character state is established, then the direction of evolution, from ancestral to derived, can be inferred, and synapomorphies can be recognized. The methodology for inferring direction of character evolution is critical to cladistic analysis. **Outgroup comparison** is the most widely used procedure. It relies on the argument that a character state found in close relatives of a group (the **outgroup**) is likely also to be the ancestral or primitive state for the group of organisms in question (the **ingroup**). Usually more than one outgroup is used in an analysis, the most important being the first or genealogically closest outgroup to the ingroup, called the **sister group**. In many cases, the primitive state for a taxon can be ambiguous. The primitive state can only be determined if the primitive states for the nearest outgroup are easy to identify and those states are the same for at least the two nearest outgroups (Maddison *et al.*, 1984).

Using the example above, determination of the primitive cetacean locomotor pattern is based on its similarity to that of the sister taxon to the cetaceans, a group of four legged mammals known as the **mesonychids** (i.e. the outgroup), which are thought to have swam by quadrupedal paddling. Locomotion in whales went through several stages. Ancestral whales (i.e. *Ambulocetus*) swam by pelvic paddling propelled by the hind limbs only. Later diverging whales (i.e. *Rodhocetus*) went through a caudal undulation stage propelled by the

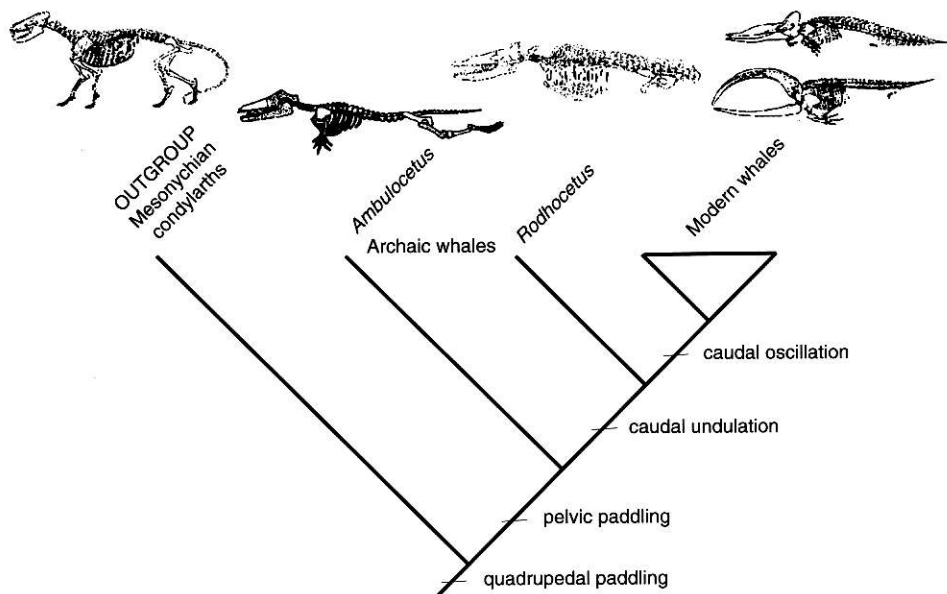


Figure 2.2 Distribution of character states for locomotor pattern among cetaceans. Skeletal reconstructions (mesonychian condylarth, *Mesonyx*, from W. B. Scott (1888); archaic whales, *Ambulocetus natans*, from Thewissen *et al.* (1996) and *Rodhocetus kasrani* from Gingerich *et al.* (1994); modern whales, an odontocete, the sperm whale *Physeter macrocephalus* and a mysticete, the bowhead whale, *Balaena mysticetus* from Oelschläger (1978)).

feet and tail. Finally, modern whales adopted caudal oscillation using vertical movements of the tail as their swimming mode (Fig. 2.2; Fish, 1993).

Derived characters are used to link **monophyletic groups**, groups of taxa that consist of a common ancestor plus all descendants of that ancestor. In contrast to a monophyletic group, a **paraphyletic group** (designated by quotation marks) is one that includes a common ancestor and some, but not all of the descendants of that ancestor. A real example of a paraphyletic group is the recognition of an extinct group of cetaceans known as “**archaeocetes**.” A rapidly improving fossil record and phylogenetic knowledge of whales now support the inclusion of “archaeocetes” as the ancestors of both baleen whales and toothed whales rather than as a separate taxonomic category (e.g., Thewissen *et al.*, 1996). It has been suggested that toothed whales may be paraphyletic, a closer relationship being suggested between sperm whales and baleen whales than between sperm whales and other toothed whales (e.g., Fig. 2.3; Milinkovitch *et al.*, 1993).

Monophyletic groups can be characterized in two ways. First, a monophyletic group can be **defined** in terms of ancestry, and second, it can be **diagnosed** in terms of characters (see Appendix 3). For example, whales or cetaceans can be defined as including the common ancestor of *Pakicetus* (an extinct whale) and all of its descendants including both modern toothed and baleen whales. Note that this definition is based on ancestry and does not change because there will always be a common ancestor for whales. On the other hand, cetaceans can be diagnosed by a number of characters (e.g. rotation of the middle ear bones and parallel arrangement of the front teeth; see also Chapter 4). The usefulness of the distinction between definition and diagnosis is that while the definition may not change, the diagnosis can be altered to reflect changes in our knowledge of the distribution of charac-

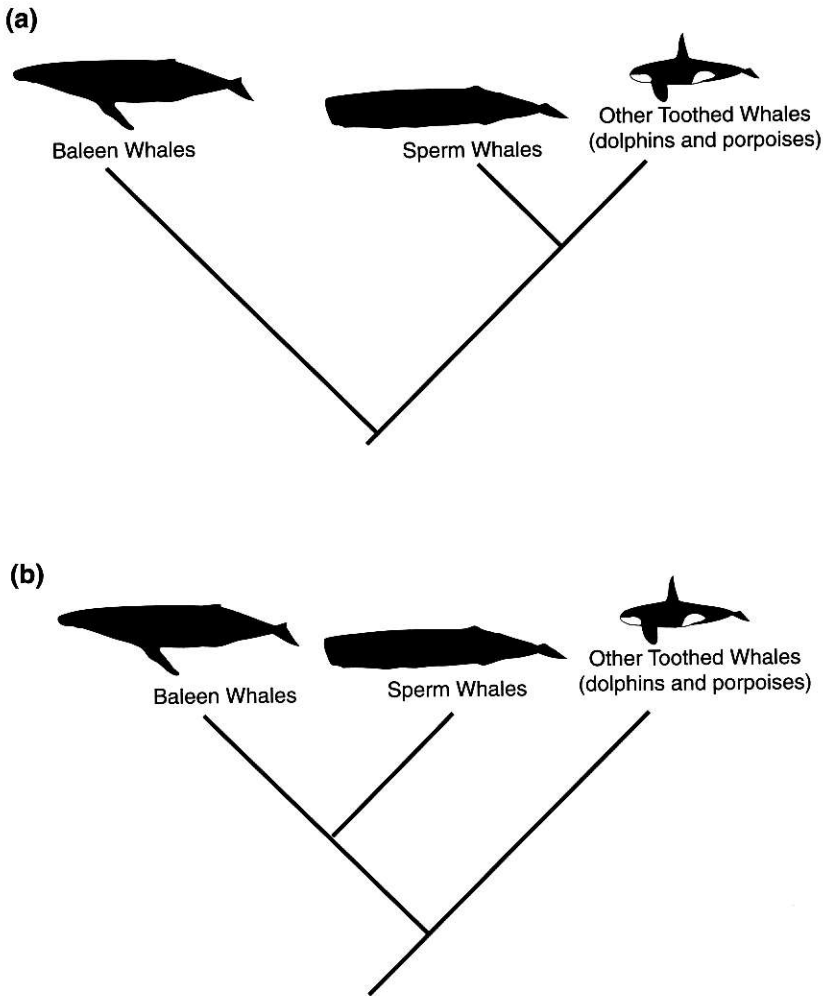


Figure 2.3. Alternative hypotheses for the phylogeny of toothed whales. (a) Traditional view of toothed whale monophyly, (b) Molecular view supporting toothed whale paraphyly.

ters. New data, new characters, or reanalysis of existing characters can modify the diagnosis. For example, in the early 1990s discoveries of new fossil cetaceans (e.g. *Ambulocetus* and *Rodhocetus*) have provided new characters illuminating the transition between whales and their closest ungulate relatives. The definition of Cetacea has not changed, but the diagnosis has been modified according to this new character information. A third term also used in this book, **characterization**, refers to a list of distinguishing features, both shared primitive and shared derived characters, that are particularly useful in field or laboratory identification of various species.

A concept critical to cladistics is that of **homology**. Homology can be defined as the similarity of features resulting from common ancestry. Two or more features are homologous if their common ancestor possessed the same feature. For example, the flipper of a seal and the flipper of a walrus are homologous as flippers because their common ancestor had flippers. In contrast to homology, similarity not due to homology is termed **homoplasy**. The

flipper of a seal and the flipper of a whale are homoplasious as flippers because their common ancestor lacked flippers. Homoplasy may arise in one of two ways: **convergence (parallelism)** or **reversal**. Convergence is the independent evolution of a similar feature in two or more lineages. Thus, seal flippers and whale flippers evolved independently as swimming appendages; their similarity is homoplasious by convergent evolution. Reversal is the loss of a derived feature coupled with the reestablishment of an ancestral feature. For example, in phocine seals (e.g. *Erignathus*, *Cystophora*, and the Phocini) the development of strong claws, lengthening of the third digit of the foot, and deemphasis of the first digit of the hand are character reversals because none of them characterize phocids ancestrally but are present in terrestrial arctoid carnivores.

It is common, but incorrect, practice to refer to taxa as being either primitive or derived. This is deceptive, because individual taxa that have diverged earlier than others may have undergone considerable evolutionary modification on their own relative to taxa that have diverged later in time. For example, otariid seals have many derived characters, although they have diverged earlier than phocid seals. In short, taxa are not primitive, although characters may be.

2.3. How Do You Do Cladistics?

Cladograms are constructed using the following steps:

1. Select a group whose evolutionary relationships interest you. Name and define all taxa for that group. Assume that the taxa are monophyletic.
2. Select and define characters and character states for each taxon.
3. Arrange the characters and their states in a data matrix (see example in Table 2.1).
4. For each character, determine which state is ancestral (primitive) and which is derived. This is done using outgroup comparison. For example, if the distribution of character #1, thick fat layers of the skin, is taken into consideration, two character states are recognized: “absent” and “present”. In Table 2.1, the outgroup (bears) have the former condition, which is equivalent to the ancestral state. This same state also is seen in one of the ingroup taxa, the fur seals and sea lions. The other ingroup taxa have thick fat layers “present” which is a synapomorphy that unites walruses and seals to the exclusion of fur seals and sea lions.

Table 2.1 Data Set for Analysis of Recent Pinnipeds Plus an Outgroup Showing Five Characters and Their Character States

Taxon	Character/Character states				
	1 Thick fat layers	2 Locomotor type	3 Pelage	4 Middle ear bones	5 Lacrimal bone
Outgroup	absent	forelimb + hind limb	abundant	small	present
Ingroup:					
Fur seals and Sea lions	absent	forelimb	abundant	small	absent
Walruses	present	hind limb	sparse	large	absent
Seals	present	hind limb	sparse	large	absent

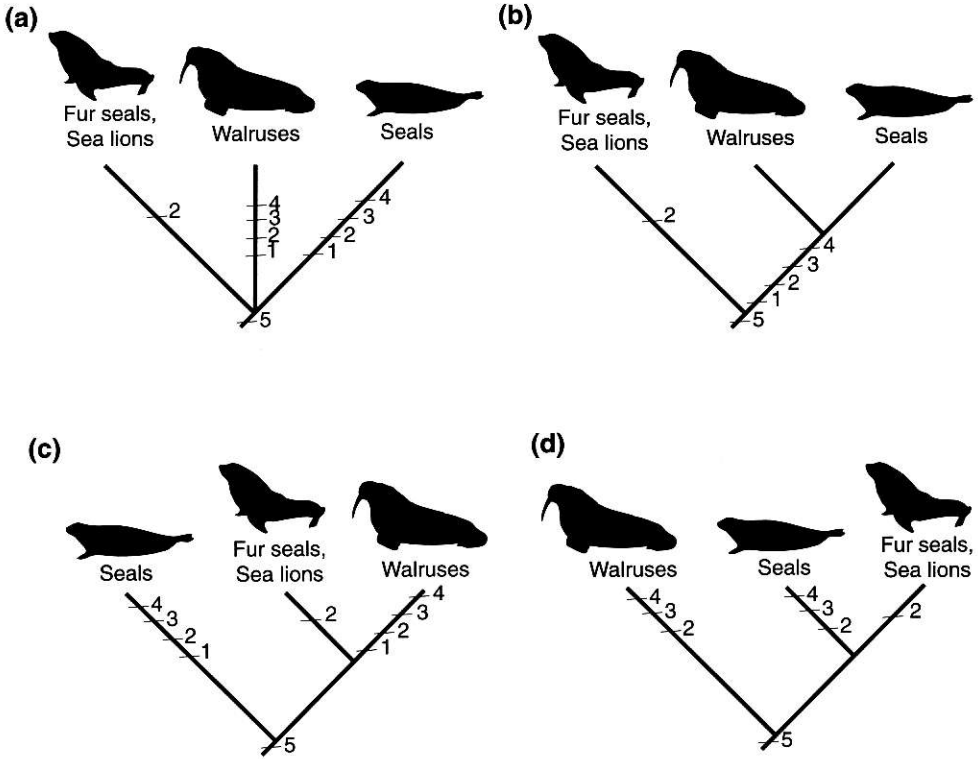


Figure 2.4. Four possible cladograms of relationship and character-state distributions for the three ingroups listed in Table 2.1. B has the most shared derived characters.

5. Construct all possible cladograms by sequentially grouping taxa based on the common possession of one or more shared derived character states (circles around character states in Table 2.1) and choose the one that has the most shared derived character states distributed among monophyletic groups (Fig. 2.4b). Note that the tree in Fig. 2.4a shows no resolution of relationships among taxa, referred to as a **polytomy**, and that the trees in Fig. 2.4c and d show mostly characters that are unique to one taxon and tell us nothing about relationships among different taxa.

2.4. Testing Phylogenetic Hypotheses

An important aspect of the reconstruction of phylogenetic relationships is known as the principle of **parsimony**. The basic tenet of the principle of parsimony is that the cladogram that contains the fewest number of evolutionary steps, or changes between character states of a given character summed for all characters, is accepted as being the best estimate of phylogeny. For example, for all the possible cladograms for the data set of Table 2.1, the one (Fig. 2.4b) illustrated in detail in Fig. 2.5a is the shortest since it contains the fewest number of evolutionary steps.

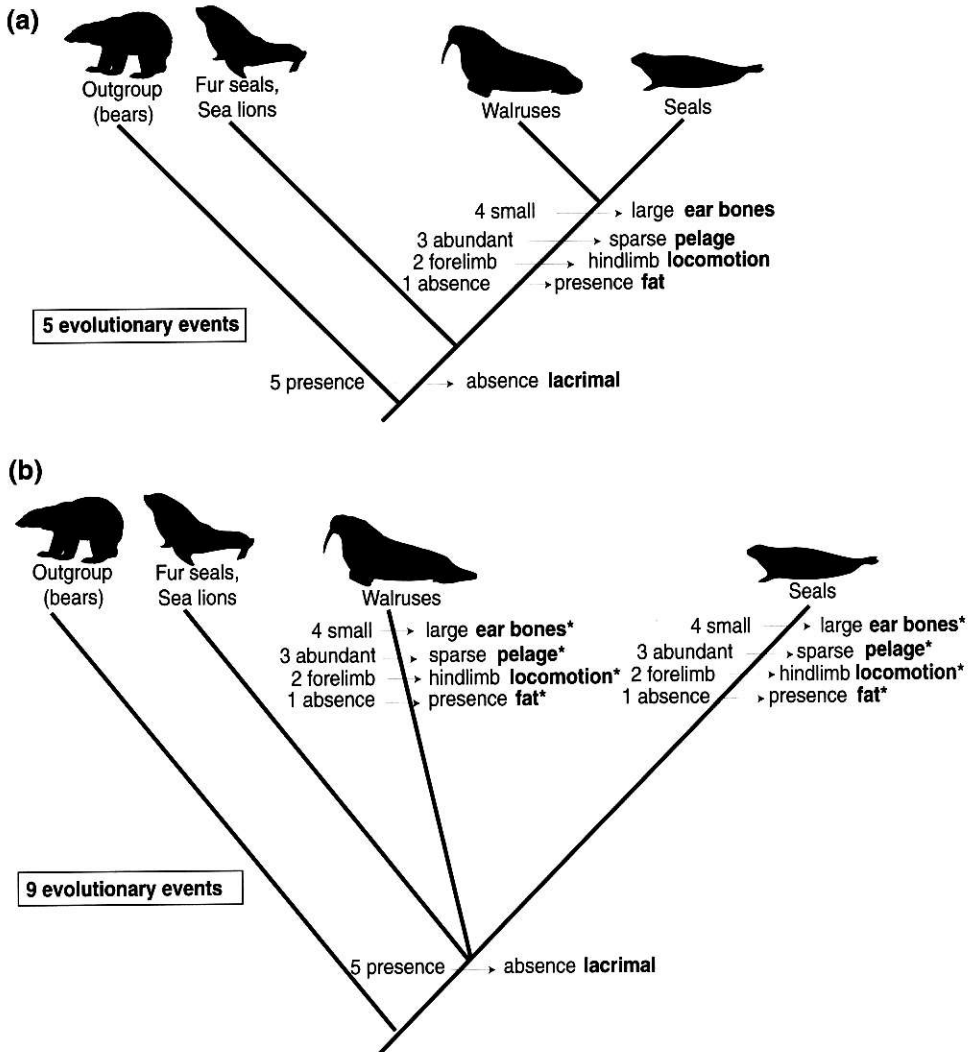


Figure 2.5. Two of the four possible cladograms. (a) Most parsimonious cladogram. Note a total of five evolutionary events. (b) Alternative cladogram showing different relationships for taxa. Note that this cladogram requires nine evolutionary events, four more than the most parsimonious cladogram. Convergences are noted by an asterisk*.

The methods used to search for the most parsimonious tree depend on the size and complexity of the data matrix. These methods are available in several computer programs [e.g. PAUP (Swofford, 1991); HENNIG86 (Farris, 1988); MacClade (Maddison and Maddison, 1992)]. The latter is particularly useful in assessing the evolution of characters. Recently, systematists have become concerned about the relative accuracy of phylogenetic trees (i.e. how much confidence can be placed in a specific phylogenetic reconstruction). Studies indicate that methods of phylogenetic analysis are most accurate if sufficient consideration is given to such parameters as sampling, rigorous analysis, and computer capabilities (Hillis, 1995).

A related issue in systematics is how to evaluate different data sets (e.g. morphology, behavior, and DNA sequences), particularly whether they should be combined (also referred to as a "total evidence" approach) or analyzed separately (Bull *et al.*, 1993). Before data sets can be combined, it is necessary to determine if they are congruent, that is the order of branching is not contradictory. This is done by what is known as the partition homogeneity test which examines data sets as partitions of characters evolving under different sets of rules. If the data sets are significantly incongruent, this suggests that the characters evolved under different models of evolution and that combined analysis would be misleading (Bull *et al.*, 1993; Farris *et al.*, 1995).

2.5. Going Beyond the Phylogenetic Framework: Elucidating Evolutionary and Ecological Patterns

Once a phylogenetic framework is produced, one of its most interesting uses is to elucidate questions which integrate evolution, behavior, and ecology. One technique used in this book to facilitate such evolutionary studies is **optimization**, or **mapping** (Funk and Brooks, 1990; Brooks and McLennan, 1991; Maddison and Maddison, 1992). Once a cladogram has been constructed, a feature or condition is selected to be examined in light of the phylogeny of the group. Examples included in this book include the evolution of body size, host-parasite associations, mating-reproductive behavior, and locomotor behavior. The condition of the terminal taxon (at the ends of branches) is identified and "mapped" onto the cladogram. There are various ways of mapping character changes onto the cladogram as discussed by Maddison and Maddison (1992). Hypothetical states are assigned to the nodes that reflect the most parsimonious arrangement of these conditions at each node. This allows one to determine the evolutionary trend of the condition in question. For example, consider the evolution of body size in phocid seals. One traditional assumption had been that small body size is the ancestral condition among phocids. This view is based on the assumption that seals of large body size represent an evolutionary advancement since they have a decreased surface area which in turn reduces body heat loss, an advantage in cold environments. This assumption, however, lacks historical evidence. When body size is mapped onto a phylogeny for seals and their relatives (walruses and sea lions; Fig. 2.6), there is a more parsimonious explanation for the data (Wyss, 1994). Accordingly, large body size is the ancestral condition for seals. A decrease in body size evolved secondarily among phocine seals (e.g. harbor, ribbon, and spotted seal). This hypothesis led Wyss (1994) to question whether this decrease in size among phocids was correlated with any other pattern of character evolution. He discovered that phocines were characterized by massive character reversals and he hypothesized that these reversals might be related to shifts in timing during development (neoteny). In addition to a decrease in body size, a number of other characters among phocines provided evidence for developmental juvenilization (i.e. failure of certain regions of the skull to ossify, resulting in perforations in the basicranium and the lack of fusion of certain cranial bones). In this example, a phylogenetic approach provided a framework for questions regarding the relationship between the evolution of body size and the pattern of evolution of other characters. A developmental explanation for the observed body size pattern was then proposed and further evidenced by other characters.

Another growing area of interest in the comparative study of phylogenies is how to deal with different types of character change, such as discrete or categorical (e.g. presence or

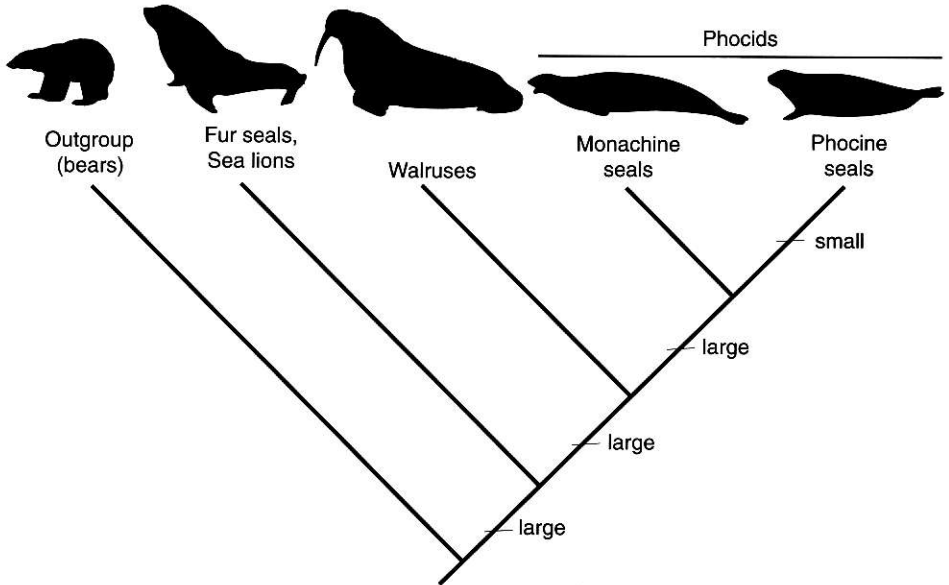


Figure 2.6. Body size mapped onto pinniped phylogeny (based on Wyss, 1994, and Bininda-Emonds and Russell, 1996).

absence of limbs) versus continuously varying characters (e.g. amount of time spent foraging). Several different methods have been proposed to incorporate phylogenetic information into comparative analyses. Examples of these techniques include Felsenstein's (1985) method of independent contrasts and the spatial autocorrelation techniques of Chevrud *et al.* (1985). These methods are designed for use with primarily continuous characters and as such are beyond the scope of this text.

2.6. Taxonomy and Classification

Classification is the arrangement of taxa (e.g. species) into some type of hierarchy. We need a system of classification so that we can more easily communicate about organisms. The taxonomic ranks are hierarchical, meaning that each rank is inclusive of all other ranks beneath it. A **taxon** is a particular taxonomic group at a given rank (plural is taxa). The major taxonomic ranks used in this book are listed below.

<i>Major taxonomic ranks</i>	<i>Example</i>
Order	Sirenia
Family	Trichechidae
Genus	<i>Trichechus</i>
Species	<i>manatus</i>

Nomenclature is the formal system of naming taxa (including the ranks of taxa) according to a standardized scheme, which for animals is the International Code of Zoological Nomenclature. These formal names are known as scientific names. The most important thing to remember about nomenclature is that all species may bear only one scientific name. The scientific name is, by convention, expressed using Latin and Greek words.

Species names are always italicized (or underlined) and always consists of two parts, the genus name (always capitalized, e.g. *Trichechus*) plus the specific epithet (e.g. *manatus*). For this reason, species names are known as binomials and this type of nomenclature is called binomial nomenclature. Species also have common names. In the previous example, *Trichechus manatus* is also known in English by its common name, West Indian manatee.

The two major ways to classify organisms are **phenetic** and **phylogenetic**. Phenetic classification is based on overall similarity of the taxa. Phylogenetic classification is that which is based on evolutionary history, or pattern of descent, which may or may not correspond to overall similarity. Phylogenetic systematists contend that classification should be based on phylogeny and should include only monophyletic groups. We have provided the most recent information on the classification and phylogeny of marine mammals. The classification of many marine mammal groups, however, is in a constant state of change due to new discoveries and information. Indeed, some systematists have offered compelling arguments for the elimination of taxonomic ranks altogether. In general, it is more important to know the names and characteristics of larger taxonomic groups like the Pinnipedia and the Sirenia than it is to memorize their rank.

2.7. Summary and Conclusions

The reconstruction of phylogenetic relationships provides a framework in biology for interpreting patterns of evolution, behavior, and ecology. Relationships are reconstructed based on shared derived similarities between species, whether similarities in morphologic characters or in molecular sequences, that provide evidence that these species share a common ancestry. The direction of evolution of a character is inferred by outgroup comparison. The best estimate (most parsimonious) of phylogeny is the one requiring the fewest number of evolutionary changes. Phylogenetically based comparative analyses have proven to be a powerful tool for generating and testing ideas about the links between behavior and ecology. The classification of species involves their arrangement in a hierarchy. Phylogenetic classification is based on evolutionary history and includes only monophyletic groups.

2.8. Further Reading

Readers are referred to texts by Eldredge and Cracraft (1980), Wiley (1981), Schoch (1986), and Wiley *et al.* (1991) for discussion of the principles and practice of phylogenetic systematics. Treatment of molecular data in phylogeny reconstruction is reviewed by Miyamoto and Cracraft (1991) and Swofford *et al.* (1996). Brooks and McLennan (1991), Harvey and Pagel (1991), Martins (1996), and Krebs and Davies (1997) provide examples of the use of phylogeny in studies of ecology and behavior.

References

- Adam, P. J. and A. Berta (1997). Osteo indicators of locomotor type and their utility in determining evolutionary patterns in the Pinnipedimorpha. *Abstr. 57th Ann. Meet. Soc. Vert. Paleo.*, Chicago, p. 28A.
- Berta, A., and T. A. Deméré (1993). The evolution of pinniped aquatic locomotor patterns. *Abstr. 10th Bienn. Conf. Biol. Mar. Mamm.*, Galveston, TX, p. 27.

- Bininda-Emonds, O. R. P. and A. P. Russell (1996). A morphological perspective on the phylogenetic relationships of the extant phocid seals (Mammalia: Carnivora: Phocidae). *Bonner Zoologische Monographien*, **41**: 1–256.
- Brooks, D. R., and D. A. McLennan (1991). "Phylogeny, Ecology, and Behavior." University of Chicago Press, Chicago.
- Brooks, D. R., R. L. Mayden, and D. A. McLennan (1992). Phylogeny and biodiversity; conserving our evolutionary legacy. *Trends Ecol. Evol.* **7**: 55–59.
- Bull, J. J., J. P. Hulsenbeck, C. W. Cunningham, D. L. Swofford, and P. J. Waddell (1993). Partitioning and combining data in phylogenetic analyses. *Syst. Biol.* **42**: 384–397.
- Chevruud, J. M., M. M. Dow, and W. Leutenegger (1985). The quantitative assessment of phylogenetic constraints in comparative analyses: Sexual dimorphism in body weights among primates. *Evolution* **39**: 1335–1351.
- Eldredge, N., and J. Cracraft (1980). "Phylogenetic Patterns and the Evolutionary Process." Columbia University Press, New York.
- Farris, J. S. (1988). "HENNIG86, Version 1.5." Distributed by the author, Port Jefferson Station, NY.
- Farris, J. S., Källersjö, M., Kluger, A. G., and Bult, C. (1995). Constructing a significance test for incongruence. *Syst. Biol.* **44**: 570–572.
- Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39**: 783–791.
- Fish, F. (1993). Influence of hydrodynamic design and propulsive mode on mammalian swimming energetics. *Aust. J. Zool.* **42**: 79–101.
- Funk, V., and D. R. Brooks (1990). Phylogenetic systematics as the basis of comparative biology. *Smithson. Contrib. Bot.* **73**: 1–45.
- Gingerich, P. D., S. Raza, Mahmood, M. Arif, M. Anwar, and X. Zhou (1994). New whale from the Eocene of Pakistan and the origin of cetacean swimming. *Nature* **368**: 844–847.
- Harvey, P., and M. D. Pagel (1991). "The Comparative Method in Evolutionary Biology." Oxford University Press, Oxford and New York.
- Hillis, D. (1995). Approaches for assessing phylogenetic accuracy. *Syst. Biol.* **44**(1): 3–16.
- Hoberg, E. P. (1992). Congruent and synchronic patterns in biogeography and speciation among seabirds, pinnipeds, and cestodes. *J. Parasitol.* **78**(4): 601–615.
- Hoberg, E. P. (1995). Historical biogeography and modes of speciation across high latitude seas of the Holarctic: Concepts for host-parasite coevolution among the Phocini (Phocidae) and Tetrabothriidae (Eucestoda). *Can. J. Zool.* **73**: 45–57.
- Krebs, J. R., and N. B. Davies (1997). "Behavioral Ecology: An Evolutionary Approach," 4th ed. Blackwell Science, London.
- Maddison, W. P., and D. R. Maddison (1992). "Mac Clade: Analysis of Phylogeny and Character Evolution, Version 3.0." Sinauer, Sunderland, MA.
- Maddison, W., M. Donoghue, and D. Maddison (1984). Outgroup analysis and parsimony. *Syst. Zool.* **33**: 83–103.
- Marsh, H., and T. Kasuya (1986). Evidence for reproductive senescence in female cetaceans. *Rep. Int. Whal. Comm., Spec. Issue* **8**: 57–74.
- Martins, E. (ed.) (1996). "Phylogenies and the Comparative Method in Animal Behavior." Oxford University Press, New York.
- Milinkovitch, M. C., G. Orti, and A. Meyer (1993). Revised phylogeny of whales suggested by mitochondrial ribosomal DNA sequences. *Nature* **361**: 346–348.
- Miyamoto, M. M., and J. Cracraft (eds.) (1991). "Phylogenetic Analysis of DNA Sequences." Oxford University Press, New York.
- Oelschläger, H. A. (1978). *Nat. Mus. Frankf.* **108**: 317–333.
- Perry, E. A., S. M. Carr, S. E. Bartlett, and W. S. Davidson (1995). A phylogenetic perspective on the evolution of reproductive behavior in pagophilic seals of the Northwest Atlantic as indicated by mitochondrial DNA sequences. *J. Mammal.* **76**(1): 22–31.
- Promislow, D. E. L. (1996). Using comparative approaches to integrate behavior and population biology. In "Phylogenies and the Comparative Methods in Animal Behavior" (E. Martins, ed.), pp. 288–323. Oxford University Press, New York.
- Reeves, R. R., B. S. Stewart, and S. Leatherwood (1992). *The Sierra Club Handbook of Seals and Sirenians*. Sierra Club Books, San Francisco.
- Reyes, J. C., J. G. Mead, and K. Van Waerebeek (1991). A new species of beaked whale *Mesoplodon peruvianus* sp. n. (Cetacea: Ziphiidae) from Peru. *Mar. Mamm. Sci.* **7**(1): 1–24.
- Reyes, J. C., K. Van Waerebeek, J. C. Cardenas, and J. L. Yanez (1995). *Mesoplodon bahamondi* sp. n. (Cetacea, Ziphiidae), a new living beaked whale from the Juan Fernandez Archipelago, Chile. *Bol. Mus. Nac. Hist. Nat. Chile* **45**: 31–44.

- Schoch, R. M. (1986). "Phylogeny Reconstruction in Paleontology." Van Nostrand-Reinhold, New York.
- Scott, W. B. (1888). On some new and little known creodonts. *J. Acad. Nat. Sci* **9**: 155–185, Philadelphia.
- Swofford, D. L. (1998). PAUP*: "Phylogenetic Analysis using Parsimony, Version 4." Sinauer Associates, Sunderland, MA.
- Swofford, D. L., G. J. Olsen, P. J. Waddell, and D. M. Hillis (1996). Phylogenetic inference. In "Molecular Systematics" (D. M. Hillis, C. Moritz, and B. Mable, eds.), 2nd ed., pp. 407–514. Sinauer Assoc., Sunderland, MA.
- Thewissen, J. G. M., S. I. Madar, and S. T. Hussain (1996). *Ambulocetus natans*, an Eocene cetacean (Mammalia) from Pakistan. *CFS. Cour. Forschungsinst. Senckenberg* **191**: 1–86.
- Wiley, E. O. (1981). "Phylogenetics: The Theory and Practice of Phylogenetic Systematics." Wiley, New York.
- Wiley, E. O., D. Siegel-Causey, D. R. Brooks, and V. A. Funk (1991). The Complete Cladist: A primer of phylogenetic procedures. *Univ. Kans. Mus. Nat. Hist., Spec. Publ.*, No. 19.
- Wyss, A. R. (1994). The evolution of body size in phocids: Some ontogenetic and phylogenetic observations. *Proc. San Diego Soc. Nat. Hist.* **29**: 69–75.

Cover photographs: Photographs on front and back covers by Pieter Folkens.

This book is printed on acid-free paper.

Copyright © 1999 by ACADEMIC PRESS

All Rights Reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Requests for permission to make copies of any part of the work should be mailed to: Permissions Department, Harcourt Brace & Company, 6277 Sea Harbor Drive, Orlando, Florida, 32887-6777

Academic Press

A Harcourt Science and Technology Company

525 B Street, Suite 1900, San Diego, California 92101-4495, USA

<http://www.academicpress.com>

Academic Press

24-28 Oval Road, London NW1 7DX, UK

<http://www.academicpress.com>

Library of Congress Catalog Card Number: 98-86074

International Standard Book Number: 0-12-093225-3

PRINTED IN THE UNITED STATES OF AMERICA

01 02 03 04 05 06 QW 9 8 7 6 5 4 3 2