

PHYLOGENETIC SYSTEMATICS

Introduction – Perhaps the best definition of Phylogenetic Systematics is found in the Journal *Systematic Biology*.

“Systematics is the study of biological diversity and its origins. It focuses on understanding evolutionary relationships among organisms, species, higher taxa, or other biological entities, such as genes, and the evolution of the properties of taxa including intrinsic traits, ecological interactions, and geographic distributions. An important part of systematics is the development of methods for various aspects of phylogenetic inference and biological nomenclature/classification.”

Next week we will explore biological diversity a bit, but today we are going to study some of the methods that **systematists** use to infer the evolutionary relationships of organisms. A brief history of the field will serve as an introduction.

For many years, no rigorous, repeatable method for classifying organisms and estimating evolutionary relationships existed. In this **traditional systematics**, researchers relied on their intuition and a lifetime of experience with a particular **taxon** to generate **classifications** and **phylogenies**. In the 1960's a new school of systematics arose, termed **phenetics** or **numerical taxonomy**. Proponents of these methods argued that organisms should be classified and phylogenies inferred based on the overall similarity of the taxa in question. Their methods were rigorous and repeatable. In the late 1960's and 1970's a new methodology, largely based on the writings of the German entomologist Willi Hennig (1966), gained standing. Proponents of **cladistics** used the principle of **parsimony** to argue that species should only be grouped using shared, derived traits (**synapomorphies**) and not shared ancestral traits (**symplesiomorphies**) or convergent traits (**homoplasies** – singular forms of all these terms end in -y). For example, body covering is a **character** whose **character states** (scales, hair, fur, feathers) can tell us a great deal about the evolutionary relationships within the Vertebrata through patterns of synapomorphies. On the other hand, the presence of a vertebral column tells us nothing about evolutionary relationships within the Vertebrata. It is a symplesiomorphy. Similarly, the presence of wings is a homoplasy in different animal groups (e.g., insects, birds, bats) and could mislead us (or pheneticists). Cladistic methods are still widely used (as are some phenetic methods) but for the analysis of DNA data, mathematically intensive statistical methods like **Maximum Likelihood** are now very popular. These methods rely on complex models of DNA and protein mutation, and prefer the phylogenetic hypothesis that is the most likely given the data.

Our Project – Working in groups of two, you will find the best phylogenies for two groups: 1) selected vertebrates and 2) eight primate species. In each case, you will employ the **principle of parsimony** – that is, you will accept the tree of shortest length (i.e., the tree hypothesizing the fewest number of character state changes) as the best estimate of the true phylogeny.

Glossary

Character – any heritable attribute of an organism.

Character state – any observed variant of a given character.

Classification – a logical system for organizing organisms.

Homoplasy – character state arising due to convergence, parallelism, or reversal.

Parsimony - preference of simpler explanations over more complex ones.

Phylogeny – a branching diagram representing the evolutionary relationships or history of genes, species, or other taxa.

Synapomorphy – shared, derived character state.

Symplesiomorphy – shared ancestral character state.

Systematist – scientist who studies systematics.

Taxon (pl. taxa) – any group of similar organisms (e.g., family, genus, species).

Taxonomy – science of classifying, defining, naming, and identifying taxa.

Remember, **synapomorphies** (shared derived character states) are evidence of common ancestry, whereas **symplesiomorphies** and **homoplasies** are **not**.

Steps to create a phylogenetic tree – Vertebrates example

1. The data files that we are going to work with today are located on the Biology 208 webpage, which can be reached via the Biology Department webpage. Copy the two data files to the desktop of your computer, either by **right-clicking on a PC**, or by **control-clicking on the Macs** and selecting **“Save Link As”** or **“Download file”** or some variation of this.
2. Next, working with your partner, open the computer program **Mesquite** (Maddison and Maddison 2004). Look for the icon with the little green tree branches. Mesquite is a Java application, and so it runs in the same fashion on Macs and PCs. (If Mesquite is not on your computer, go to Mesquiteproject.org to download it.) If any error or other messages appear, click the **“OK”** button or its equivalent. Once the program is loaded, go to the **“File”** menu and click **“Open File”**. Locate the desktop and open the file named **“Vertebrates.nex.txt”**. The **“.nex”** suffix means that this file is in Nexus format, which is a commonly used by phylogenetic software. The **.txt** suffix was added to the file so that it could be transmitted via the WWW.
3. A navigation window will open. Click on the button labeled **“Show Matrix”**. A pane will open that shows a morphological data matrix for a select group of vertebrates. The taxa are listed down the left side, and the characters are listed across the top. Maximize the window so that you can see all the character states. Do some vertebrates appear to be related with one another because of shared character states? Based on this cursory examination, make a few hypotheses about evolutionary relationships or clades of taxa. Remember to differentiate between shared, ancestral character states and shared, derived character states.
4. Next, you will rigorously test your hypotheses using Mesquite. This software provides an interactive environment for exploring phylogenetic hypotheses. It will allow you to quickly manipulate hypothesized phylogenetic trees to search for the most **parsimonious tree**. As trees are manipulated, the program updates statistics such as **tree length** and the results are illustrated graphically.
5. Pull down the **“Taxa&Trees”** menu and select the first option, **“New Tree Window.”** For the following window, simply select the **top** or default **option (“Stored Trees”)** by clicking **“OK”** (do not select the default trees). A new pane will open, this time displaying a phylogenetic hypothesis for the vertebrates. This tree was stored in the data file with the character state data.
6. We need to know how many evolutionary steps are on our tree. So, in the tree window, pull down the **“Analysis”** menu. Select the **“Values for Current Tree”** option. In the following popup window, select the default or top option. A legend will appear with the name of the tree (Bad Guess) and its length (35 steps).
7. In order to guide your rearrangement of branches, you can map character state changes onto the tree to identify synapomorphies. Click again on the **“Analysis”** window. Select **“Trace Character History.”** Again, choose the default, or top options for the next popup. Now a window appears showing the state changes in character 1 mapped onto the tree. If you click on the **blue arrow**, you can see the changes for the other characters. Use them

to guide your arrangement of branches, trying to group together taxa with the same character states.

8. Now, as you click on a branch and hold it, you can drag it to a new position within the tree. Once you release the branch, watch the tree length window. Accept only those moves that reduce the overall tree length (that make it more parsimonious). You will move (swap) branches until you find the shortest tree (the tree with the fewest number of steps). You can use the “**Edit**” menu or **control-z** (**apple-z** on Macs) to undo a move. Use the arrow icons to step through each character and alter the groupings until you can no longer reduce the tree length. What’s the most parsimonious tree for these selected vertebrates?
9. Map character 4 onto your tree. What is going on here? **You may want to note answers to question 3 below** before you close the vertebrates file.
10. There is no simple way to save a picture of the tree in Mesquite. In order to do this for your lab write-up, the best way is to use “**File: Save Tree as PDF.**” You can then copy it from the PDF using the **Snapshot Tool** and paste it into MSWord for you lab report.

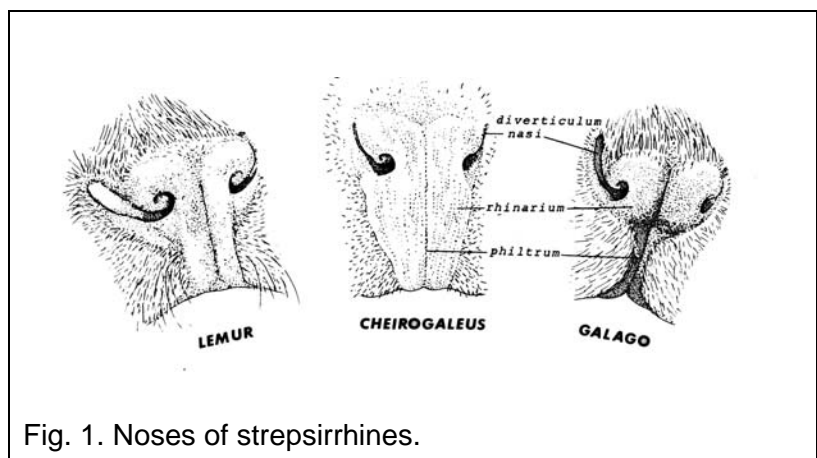
Finding the most parsimonious phylogeny for eight primate species

In this exercise, you will examine primate skull replicas and decide how to score the various characters for each species. Some of the characters required for this analysis are post-cranial and ecological. You will find the states of these characters in the **following introductory paragraphs**. We will introduce you to the skull characters during lab, after you have had a chance to examine the skulls for a few minutes on your own.

The primates include the Suborders **Strepsirrhini** and **Haplorrhini**. With a few exceptions (e.g., us), primates all have grasping hands and feet. Hindlimbs dominate locomotion, and eye orbits are rotated forward to allow stereoscopic vision (Patton 2005). **All** of the primates below have **two incisors and one canine tooth** on each side of their upper and lower jaws. The teeth beyond these, known as **cheek teeth**, vary in number. **Be careful** as you count them! Sometimes the first premolar can look like a canine tooth. With the exception of the great apes, all primates have **tails**.

The strepsirrhines, also known as prosimians, are an **Old World** group that includes lorises, pottos, bushbabies, and lemurs.

We are considering them the outgroup in today’s lab. In addition to several traits visible on their skulls, this group can be recognized by their **laterally-directed, slit-like nostrils** and the presence of **rhinarium**, a patch of bare skin adjacent to the nostrils (think of a wet dog nose; Fig. 1; Patton 2005).



Two major groups of haplorrhines are generally recognized. The **Platyrrhini**, also known as the **New World monkeys**, include howler, capuchin, spider, squirrel, owl, and other monkeys. The only primates with **prehensile tails** (tails that the monkey can hang from) are New World Monkeys (e.g., capuchins and howler monkeys). The nostrils of **platyrrhines are simple (not slit-like), widely separated, and directed laterally** (Fig. 2, top).

The **Catarrhini**, evolved in Africa and Asia, include Old World monkeys (e.g., macaques, baboons) and great apes (e.g., chimpanzees, apes, humans). Their (our) **nostrils are simple, close together, and directed forward**

(Fig. 2, bottom; Patton 2005). With the exception of the family Hominidae (i.e., the great apes, including humans and chimps), the catarrhines have **ischial callosities**, or leathery patches of bare skin on either side of their rumps (Great Ape Trust 2007). The New World monkeys do not have ischial callosities. No haplorrhines have **rhinaria**.

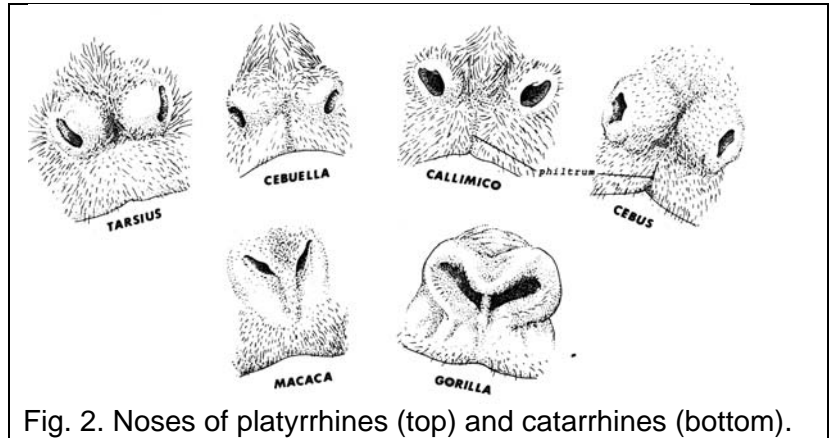


Fig. 2. Noses of platyrrhines (top) and catarrhines (bottom).

The primates that we will be studying include the following species.

The potto, *Perodicticus potto*, is a member of the Loris family. These prosimians live in the canopy of rainforests in western, tropical, sub-Saharan Africa. They are horribly cute, nocturnal and feed on fruit and insects (Skulls 2007).



The potto.

The ring-tailed lemur, *Lemur catta*, is a prosimian from southern Madagascar. They are diurnal and omnivores, eating fruit, leaves, and insects (Skulls 2007).

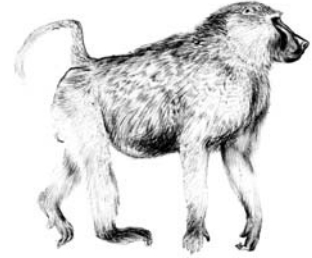
The Mexican Howler Monkey, *Alouatta pigra*, is arguably one of the loudest animals in the world. These monkeys mainly eat leaves and other plant material, though the occasional invertebrate may be consumed as well. They can be found in rainforests from southern Mexico into South America (Skulls 2007).



The brown capuchin (relative of weeping).

The weeping capuchin, *Cebus olivaceus*, inhabits South American rainforests. This medium sized monkey is an omnivore (Skulls 2007). It is a highly intelligent species, named after the Capuchin friars that it resembles. Capuchins have been labeled “organ grinder monkeys” in popular culture (Wikipedia 2007a). They are closely related to the brown capuchins in the Bucknell primate colonies.

The hamadryas baboon, *Papio hamadryas*, is a species found in the Bucknell primate colonies. This is the smallest of the baboons and is native to East Africa (Skulls 2007).



The hamadryas baboon.

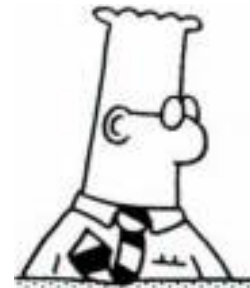


The rhesus macaque. (Wikipedia 2007b).

The rhesus monkey, *Macaca mulatta*, inhabits a range covering much of southern China and stretching from Afghanistan to India and Nepal. This omnivorous species is highly adaptable, living in habitats ranging from deserts and snowcapped mountains to human cities (Skulls 2007).

Chimpanzees, *Pan troglodytes*, are native to central and west Africa. They are highly social omnivores. Like the other great apes, they do not have tails. They have been known to live as long as 74 years in captivity

Humans, *Homo sapiens*, are the most widely distributed primates of all, with constant representation in nearly every conceivable habitat, including all continents, tens of thousands of islands, the deep sea, underground to depths of several miles, trees, the atmosphere to elevations of about 50,000 feet, and outer space. Genetic evidence, however, strongly suggests that our species evolved in East Africa. We are omnivores, and, with the exception of EE majors playing video games all night, most comfortably diurnal.



Nocturnal *Homo sapiens*.

- In our exercise today, we consider the strepsirrhines to be the **outgroup** and haplorrhines to be the **ingroup**. The ingroup is the taxon of primary interest in a phylogenetic study. Outgroups can be very useful in such studies, and are included in most. They are used to root the phylogeny (polarize it temporally). The character states of the outgroup are often assumed to be ancestral to those of the ingroup, and are used to infer the direction of evolution.

Steps to create a phylogenetic tree – Primate example

1. Examine the eight primate skulls, looking for characters that might be phylogenetically informative. Use the general look and feel of the skulls to intuitively sketch a phylogeny for them. Try to ignore any prior information that you might have about the primate as you do this, focusing on the data in front of you. Do any characters look particularly promising? Spend at least ten minutes on this step.
2. Ask the instructors for the summary sheet that will introduce you to primate skull morphology. After this orientation, **use the blank worksheet at the end of this lab** to score the states of 14 selected primate characters. Cranial characters 1-8 should be scored based on the skulls. **All other characters are contained in the descriptions above.**
3. Next, rigorously test your hypotheses using the computer software **Mesquite**. First, use the “**File**” menu to close the Vertebrates data file.
4. Begin by opening the “**Primates_worksheet.nex.txt**” file in the data editor window. To save you time, we have already entered the names and states for each character as well as

the taxon names. You will need to replace the question marks with the appropriate character states.

5. Enter the states based on the values in Table 1. Hint: If you type the number that corresponds to the character state you want, Mesquite will fill in the appropriate text for that character state. Once your data are entered, pull down the “**Taxa&Trees**” menu and select “**New Tree Window.**” You should select “**Default Trees**” to represent the evolutionary relationships among the primate species. Three default trees will be opened. Click on the **blue forward arrow** in the upper left-hand corner of the window until the tree labeled “**Default Ladder**” is showing (it may be already). **Follow the same steps as above to view the tree length and to trace the characters on the tree.** Tracing the characters will allow you to quickly visualize the species that share character states. Again, remember that it’s the shared, derived states that are important in testing your hypotheses about phylogenetic relationships.
6. Work through each character and try to group those species that **share derived** states. **Watch the tree length** window to see the change to tree length – accept only those manipulations that shorten the tree (why do we do this?). You can **undo manipulations** by using the “**Edit**” menu or by typing “**control-Z**” or “**apple-Z.**”
7. Continue until you’re unable to shorten the tree length. **Root it using the outgroup** (the strepsirrhines). To do this, simply attach the branch that leads to the outgroup taxa to the base of the tree. Starting at the bottom of the tree, the first bifurcation should lead to the outgroup on one side and the other six species on the other side. **(If you do not root the tree properly, you may miss some important points below. Ask for help if you need it.)** This should not change the tree length. If it does, adjust the other groups to return to the shortest tree. Does the phylogenetic tree that you’ve generated make sense? How does it compare to the hypotheses that you and your partner made before rigorously testing? What is going on with character 6, the simian shelf?
8. Save a picture of your tree (see above), with **Character 6** traced on it, for inclusion in the write-up.

Your Write-up – At next week’s lab, hand in a typewritten, single-spaced summary of your findings that includes REDUCED copies of the two trees that you generated, one for **vertebrates** and one for the **primates**.

Answer the following questions.

1. What was the shortest tree length that you could generate for Vertebrates and for the Primates? What’s the usefulness of tracking tree lengths (i.e., why do we use this measure)?
2. Use a technical systematics term defined in the introduction above to describe the evolution of character 4 in the vertebrate data set and character 6 in the primate data set. What evolutionary phenomenon can lead to the pattern generated when these characters are mapped onto the phylogenetic tree? If you changed the topology of the tree so that this pattern vanished, what would happen to the tree length? Why? Use a different term to describe the evolution of character 5 in the primate data set. If this character was not in the data set, would you have found the same shortest tree?

3. Which synapomorphies in the dataset unite birds and crocodiles? (Be careful here – this means character states shared **only** by birds and crocs!) Give an example of one other synapomorphy from the vertebrate data set and another from the primate data set. (There are, of course, many to choose from.) List the taxa united by these characters. Give an example of a character that would be a symplesiomorphy for the vertebrate data set (use your knowledge of vertebrates – there are no symplesiomorphies in the data set).
4. Recall your pre-tree inspection of each data set. Were you able to correctly predict which groups of taxa would cluster together before you actually manipulated the tree? Why or why not? Would your success change if there were 50 taxa and 1500 characters (a smallish data set by today's standards)? Are you 100% that the inferred primate relationships are true? Why or why not? How could you further test them?
5. Which primate species are most closely related to the outgroup?
6. In a short paragraph, discuss several qualities of a good phylogenetic character and several qualities of a bad phylogenetic character.
7. List one pro and one con for three of the phylogenetic methods or schools presented in lecture.
8. Were the character states of the skulls easy to ascertain in all cases? What challenges do all schools of systematics have when it comes to coding characters?
9. There are many possible systems for classifying organisms. We could, for example, organize them based on color or alphabetical order. These systems would be next to useless, of course. If you wanted to follow in the footsteps of Linnaeus by developing a system for classifying organisms, would you want the taxonomic categories (e.g., families, orders, genera) to be monophyletic? Why or why not? (Note – there is no obviously correct answer to this question. It is designed to encourage you to think carefully about the issues involved in a successful classification, including our natural biases as humans.)

Acknowledgements

Many thanks to Dr. DeeAnn Reeder and Dr. Tristan Stayton for patiently working out the primate characters with us. Characters were derived from Patton (2005) and Whitehead et al. (2005).

Literature Cited

- Great Ape Trust. 2007. Great Ape Trust Glossary. Retrieved Feb. 2, 2007.
<http://www.greatapetrust.org/gatlibrary/glossary.php>
- Hennig, W. 1966. Phylogenetic Systematics. University of Illinois Press, Urbana.
- Maddison, W. P. and D.R. Maddison. 2004. Mesquite: a modular system for evolutionary analysis. Version 1.05 <http://mesquiteproject.org>
- Patton, J. L. 2005. Orders and Families of Living Mammals: A Lab Manual for Mammalogy. Museum of Comparative Zoology, UC Berkeley.
- Skulls Unlimited. 2007. Primates and links therein. Retrieved Feb. 2, 2007.
<http://www.skullsunlimited.com/primates.html>
- Whitehead, P. F., W. K. Sacco, and S. B. Hochgraf. 2005. A Photographic Atlas for Physical Anthropology. Morton Publishing Company, Englewood, CO.
- Wikipedia. 2007a. Capuchin monkey. Retrieved Feb. 2, 2007.
http://en.wikipedia.org/wiki/Capuchin_monkey
- Wikipedia. 2007b. Chimpanzee. Retrieved Feb. 2, 2007. <http://en.wikipedia.org/wiki/Chimpanzee>

Table 1. Characters and character state codes for species of primates. The number entered in each square for the character state is appropriate for use in character coding in Mesquite.

Common name	Species	Family	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Mexican howler monkey	<i>Alouatta pigra</i>	Cebidae														
Rhesus macaque	<i>Macaca mulatta</i>	Cercopithecidae														
Potto	<i>Perodicticus potto</i>	Lorisidae														
Human	<i>Homo sapiens</i>	Hominidae														
Weeping capuchin	<i>Cebus olivaceus</i>	Cebidae														
Hamadryas baboon	<i>Papio hamadryas</i>	Cercopithecidae														
Ring tailed lemur	<i>Lemur catta</i>	Lemuridae														
Chimpanzee	<i>Pan troglodytes</i>	Hominidae														

Characters:

1. post-orbital plate – 0 if absent, 1 if present
2. tooth comb – 0 if absent, 1 if present
3. orientation of foramen magnum– 0 if posterior, 1 if ventral
4. number of cheek teeth
5. number of incisors
6. simian shelf – 0 if absent, 1 if present
7. width of brain case – 0 if <10 cm, 1 if >10 cm
8. auditory tube – 0 if absent, 1 if present
9. tail – 0 if absent, 1 if non-prehensile, 2 if prehensile
10. ischial callosities – 0 if absent, 1 if present
11. rhinarium – 0 if absent, 1 if present
12. orientation of nostrils – 0 if lateral, 1 if forward
13. nostrils – 0 if slit-like, 1 if not slit-like
14. native range – 0 if Old World, 1 if New World