

# Evolution and Biodiversity Laboratory

## Systematics and Taxonomy

by Dana Krempels and Julian Lee

Recent estimates of our planet's biological diversity suggest that the species number between 5 and 50 million, or even more. To effectively study the myriad organisms that inhabit the biosphere, we attempt to classify organisms into groups that reflect evolutionary relationships.

### I. Taxonomy

Strictly speaking, **taxonomy** is the science of sorting and classifying living organisms into groups called **taxa** (singular = **taxon**). Taxonomy also includes describing and naming the members of those taxa. A scientist who engages in taxonomy is a **taxonomist**.

A taxon is a group of organisms that a taxonomist has judged to represent a cohesive unit. The criteria used to sort specimens into various taxa are not fixed, and the science of taxonomy is not without its internal controversies.

Taxonomists often distinguish between **natural** and **artificial** taxa. A natural taxon is constructed on the basis of evolutionary relationships. While not all taxonomists insist that taxa be natural, most believe that taxonomic groups should consist of evolutionarily related units. The science of determining evolutionary relationships among taxa is known as **systematics**, and its practitioners are **systematists**. (Most systematists are also taxonomists, and vice versa.)

Since systematists are concerned not only with the ability to sort and identify organisms, but also with determining their evolutionary relationships, taxonomy is used as a tool within systematics.

Biological **nomenclature** is the application of names to organisms recognized as part of a particular taxon. From most to least inclusive, the major taxonomic **ranks** are shown in Figure 1.

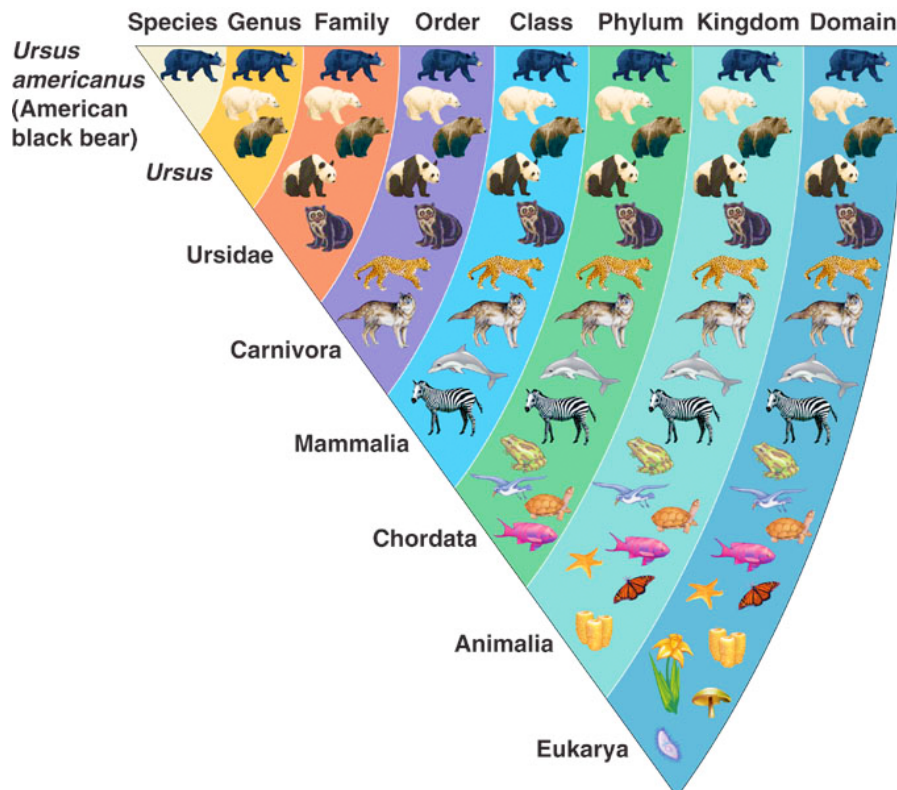


Figure 1. The Linnaean taxonomic hierarchy.  
systematics-1

Each **Domain** contains related **Kingdoms**. Each kingdom consists of related **phyla**. Each phylum consists of related **classes**, classes of related **orders**, orders of related **families**, families of related **genera** (singular: genus) and genera of related **species**. ("King Philip came over from Germany stoned.")

Between the major taxonomic ranks may be larger and smaller taxa such as **subkingdoms**, **superphyla**, **subclasses**, **infraorders**, **subspecies**, etc.

Every described, named organism is nested into a complete organizational hierarchy, from species through domain.

**The scientific name of an organism (its genus and species) is always written with the genus capitalized and the specific epithet in lower case. Because the words are Latinized, they should be italicized.**

This system of nomenclature was created by Swedish botanist Carl Linne, who published it as *Systema naturae*, in 1735. Linne Latinized his own name to Carolus Linnaeus, and we remember him today as **Linnaeus**, the father of modern taxonomy.

## **A. The Aspects of a Taxon**

A taxon is generally considered to have three aspects:

### **1. The taxon's name.**

The name of the taxon to which all flesh-eating mammals with specialized cutting teeth called carnassials belong is **Carnivora**. The name of the taxon containing all domestic dogs is *Canis familiaris*. You get the idea.

A scientific name has no more significance than any other convenient label used to describe a group of similar items. Names such as "Bacteria," "Felidae" and "*Oryctolagus cuniculus*" are similar in function to descriptive names of similar objects, such as "shoes" or "machines."

Don't let names confuse or intimidate. Once you know the Latin or Greek word roots, seemingly complicated names make perfect sense and become easier to remember. For example, the name of *Eleutherodactylus planirostris*, a little frog naturalized in southern Florida gardens, can be broken down into its Greek roots: *eleuthero*, meaning "free," *dactyl*, meaning "toe," *plani*, meaning "flat" and *rostris*, meaning "nose." Our pal is just a flat-nosed frog with unwebbed ("free") toes!

### **2. The taxon's rank.**

Like the taxon's name, the taxon's rank has no real biological significance. It serves only to help the biologist locate the taxon within its hierarchy. For example, the taxon "**Eukarya**" is currently assigned the rank of **domain**. The taxon "**Mammalia**" is currently assigned the taxonomic rank of **class**.

**A taxon's rank can change.** You may notice that a given taxon's rank may not always be the same in every source you read. Some publications may refer to Basidiomycota (Club Fungi) as a phylum, whereas others might refer to it as a subphylum. Classifications change as new data become available.

The *relative* rank of a taxon within its larger and smaller groupings is more relevant than the rank itself. For example, it's important to know that all members of *Felis* are classified within the larger taxon "Carnivora," and that all carnivores are classified within the still larger taxon "Mammalia." It's *less* important to struggle to recall that "Carnivora" is an order and "Mammalia," is a class.

Many institutions use a **rankless system**. In this system, a taxon is described only by its name. The rank is left off, but tacitly understood. An author using this system will write "Mammalia" rather than "Class Mammalia" avoiding confusion if its rank changes.

### 3. The taxon's content.

All the students in your lab are (probably) members of the genus *Homo* and the species *Homo sapiens*. To the systematist, this is perhaps the most relevant aspect of the taxon. By grouping individuals within a single species, related species within a single genus, related genera within a single family and so on, the systematist tells us which organisms share common evolutionary ancestry.

Organisms are not classified randomly. The systematist uses morphological characters, DNA sequencing, protein analysis, developmental biology, karyology, ultrastructure and other information to determine evolutionary relationships.

## B. The Taxonomic Key: A Tool for Identification

When an investigator must identify an unknown specimen, a useful tool is the **taxonomic key**. A taxonomic key is constructed as a series of paired statements/descriptions based on similarities and differences between taxa in a group being identified. Because the key branches in two at each stage, is called a **dichotomous** (from the Greek *dicho* meaning "in two" or "split" and *tom*, meaning "cut") **key**.

Paired statements describe contrasting characteristics found in the organisms being classified. With the specimen at hand, the investigator chooses which of the paired statements best matches the organism. The statement selected may immediately identify the specimen, but more often it will direct the user to the next set of paired, descriptive statements. At the end--if an appropriate key has been used (e.g., you wouldn't use a book called *Key to the Flora of Southern California* to identify an unknown tree you've discovered in Guatemala)--the specimen is identified by name.

Sometimes a key for identification of a specimen you have at hand simply doesn't exist, and you must go to the primary literature to see if any species descriptions match it. Identification of unknown species can be a difficult and challenging enterprise. Fortunately, the specimens you're going to use in today's first exercise are not only easily recognizable, but also included in a ready-made key.

## Exercise I. Using a Taxonomic Key

**Work in pairs for this exercise.** At your station you will find several containers filled with "species" of pasta native to the United Aisles of Publix. The noodles have an evolutionary relationship to one another: They all are members **Order Semolina**, which evolved from an ancestor resembling a soda cracker. The taxonomic key—which may or may not reflect their evolutionary relationships—is a tool that allows identification of an individual pasta to its proper taxonomic group. In this case, the key identifies each type of pasta to genus and species.

Let's **key out** (a jargon-y verb commonly used to describe the process of identifying things with a taxonomic key) some pasta! Select one individual from each of the containers, place it in one of the plastic cups provided, and then use the taxonomic key below to identify each pasta individual to its correct genus and species.

**A NOTE OF CAUTION:** Be careful when choosing which of the two character states in the key your pasta actually exhibits. What exactly is its skin? What is its body? What is its body form? Confusing traits can cause incorrect identification.

This is true in real keys used to identify real organisms, too. Character states are not always obvious, and some types of organisms (Chenopode plants!) are notoriously difficult to identify, even with an excellent taxonomic key. So proceed with caution, and if you do make an error, go back and start from the beginning.

# A TAXONOMIC KEY TO THE PASTA OF SOUTHERN FLORIDA

- 1a. Body tubular in shape ..... 2
- 1b. Body not tubular ..... 4
  
- 2a. Skin lined with small, symmetrical ridges ..... 3
- 2b. Skin smooth ..... *Ziti edulis*
  
- 3a. Anterior and posterior ends of organism slanted ..... *Penna rigata*
- 3b. Anterior and posterior ends of organism  
perpendicular to body axis ..... *Rigatonii deliciosus*
  
- 4a. Skin lined with small, symmetrical ridges ..... *Conchus crispus*
- 4b. Skin not lined with ridges ..... 5
  
- 5a. Body cylindrical in overall shape ..... *Rotinii spiralis*
- 5b. Body dorsoventrally flattened in shape ..... *Farfalla aurea*

**Write the name of each type of pasta underneath its picture below.**



name:



name:



name:



name:



name:



name:

## Exercise II. Creating a Taxonomic Key

**Work in pairs for this exercise.** Now that you have seen how simple it is to use a taxonomic key, you should be able to create one. At your station you will find a set of eight cards bearing pictures of imaginary animals. These hypothetical animals, created and "evolved" by **J. H. Camin**, Professor of Biology at the University of Kansas, are called **Caminalcules**. (An "animalcule" is a small animal). Caminalcules have served as test material for a number of experiments in systematic theory and practice. Use of imaginary organisms for such studies offers a distinct advantage over using real groups, because preconceived notions and biases about classifications and evolutionary relationships can be eliminated.

Create a dichotomous key of your Caminalcule species (omit the OUTGROUP, on the light gray card; use only the numbered Caminalcules). Refer to the pasta key from the previous exercise to guide your organization. There's no single correct way to create a taxonomic key. The one you used to identify your pasta "species" could have been arranged in many other ways. It is not required that a key reflect evolutionary relationships, though many keys do. Once you have completed the second part of today's lab (Systematics), you'll be better prepared to create a key that reflects common ancestry. But for now, it's not necessary.

Use your paperback copy of *A Guide to Greek and Latin Word Roots* by Donald J. Borror to create a Latinized scientific name (consisting of genus and species) for each of your species, and try to be as descriptive as possible with the name. (Some of your individuals might be in the same genus. It's for you to decide.) Use proper *Systema naturae* rules in naming your species: Genus capitalized, species lower case, and name italicized. (If you don't have a copy of the Borror book, you may get a loaner from your TA, in exchange for your Cane Card. No Cane Card, no loaner.)

### A Taxonomic Key for Identification of Caminalcules

1a. \_\_\_\_\_

1b. \_\_\_\_\_

2a. \_\_\_\_\_

2b. \_\_\_\_\_

3a. \_\_\_\_\_

3b. \_\_\_\_\_

4a. \_\_\_\_\_

4b. \_\_\_\_\_

5a. \_\_\_\_\_

5b. \_\_\_\_\_

6a. \_\_\_\_\_

6b. \_\_\_\_\_

7a. \_\_\_\_\_

7b. \_\_\_\_\_

Once you have finished your key, trade it AND the cards used to devise it with the lab partners across the table from you. (Each team has a different set of Caminalcules, so you'll need to use the other team's cards, too.) Using each other's keys, try to identify all of each other's species correctly. When you have identified them all, check with your "swap buddies" to see how well you did.

## II. Systematics

Because new data constantly change our understanding of evolutionary relationships, classifications are constantly updated and changed. The goal of most modern systematists is to construct **monophyletic** taxa, which reflect true evolutionary relationships by including all descendants of a single common ancestor. Various lines of evidence can be used to determine the degree of common ancestry between two taxa, including comparison of morphology (at many levels, including cellular), nucleic acid sequence, protein sequence, embryo development, etc. As new technologies arise, our ability to study evolutionary relationships evolves.

### A. Reconstructing Phylogenies

A **phylogeny** is a history of the evolutionary descent of **extant** (i.e., presently living) or **extinct** (i.e., no longer living) taxa from ancestral forms. To date, about 1.4 million species (including 750,000 insects, 250,000 plants and 41,000 vertebrates) of the 5 to 50 million on earth have been scientifically described and classified.

What is a **species**? Although biologists still debate the precise definition, we shall use the biological definition of a species as **a group of actually or potentially interbreeding natural populations which are reproductively isolated from other such groups**. More simply, two organisms can be considered members of the same species if they can breed to produce fertile, viable offspring under natural conditions.

#### 1. Primitive vs. Derived Characters

Ever since Darwin's publication of *On the Origin of Species by Means of Natural Selection*, the scientific community has labored to understand how different species arise. We know that extant species **evolved** from previously existing **ancestral species**, and that this may involve **descent with modification** of **traits** (= **characters**) from one generation to the next. Terminology:

- **primitive character** (**plesiomorphy**) shows little or no change from the same character in an ancestor
- **symplesiomorphy** (literally "shared primitive character") is a primitive character shared between two or more taxa
- **derived character** (**apomorphy**) has changed in appearance and/or function relative to the same character in an ancestor
- **synapomorphy** (literally "shared derived character") is a derived character shared between two or more taxa

#### All living things exhibit these most basic symplesiomorphies:

1. Organization of structure (**anatomy**)
2. Capacity to generate more organisms like themselves (**reproduction**)
3. **Growth** and development
4. Ability to utilize energy to do work (**metabolism**)
5. Response to environmental stimuli (**reaction**)
6. Regulatory mechanisms to keep the internal environment within tolerable limits (**homeostasis**)
7. Populations that change in gene composition over time (**evolution**)

Consider: Would your knowing only that a living thing has the ability to maintain homeostasis help you distinguish it from other living things? Would knowing only that it could reproduce allow you to tell it apart from other living things? Simple answer: NO. Shared, primitive characters are not informative to someone trying to sort the organisms into smaller, less inclusive groups.

In classifying members of a taxon, the systematist must consider characters that make the individuals in that taxon *unique* and *different* from members of other taxa. To achieve this end, synapomorphies unique to that taxon are informative and useful. The next section explains why.

## 2. **Symplesiomorphies vs. Synapomorphies**

Because all living things share evolutionary history, however distantly, each taxon shares certain very ancient (i.e., **primitive**, or **plesiomorphic**) characters with other taxa. Shared, primitive characters cannot be used to separate members of different taxa, since everyone has them. However, more recently evolved (i.e., **derived**, or **apomorphic**) characters *can* set one taxon apart from another. Synapomorphies inherited from a common ancestor can inform the systematist about relative **recency of common descent**. [The more synapomorphies two taxa have in common, the more recently they evolved from a common ancestor.](#)

We humans share certain characters, unique to animals, with *all* other animals but not with plants, fungi, protists, or bacteria. List five characters unique to humans and all other animals, *but not found in any other living things* (e.g., plants, fungi):

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_
4. \_\_\_\_\_
5. \_\_\_\_\_

**Important:** the characters you listed above--exhibited by no living organisms except animals--are considered *symplesiomorphies only with respect to Animalia*. But if you include all living things, then these same animal characters become *synapomorphies that set animals apart from all other living things*. A character cannot be "primitive" or "derived" in a vacuum. It can be described with these terms only when taxa and their characters are being compared.

With this in mind, list three derived characters that set mammals apart from all other animals:

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_

Do you exhibit all three of the characters listed? (Good! You're a mammal!) Since you share those characters with all your mammalian relatives, the characters are said to be **primitive with respect to all mammals**, though they **are derived with respect to all animals other than mammals**.

See the pattern? Because you share the three characters above with all other mammals, those characters won't help you determine how closely related you are to any other mammal groups. Hence, we must consider synapomorphies at the next level, to separate our taxonomic group within the rest of the mammals.

List three derived characteristics shared by all primates (Primates, of which you are a member), but not shared by other mammals. (You might have to do some searching.)

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_

What you have listed are three synapomorphies shared by Primates that set them apart from all other mammals. But because all primates share these three characters, they are symplesiomorphies with respect to only primates. In other words, these three characters will not help you to determine which primates are your closest relatives. To do that, we must find more unique derived characters.

List two derived characteristics shared by all great apes (Hominidae, of which you are a member), but not shared by other primates. Again, you might have to do some searching. Notice that it can become more difficult to find synapomorphies linking particular members into a single as the taxon becomes smaller/less inclusive, because organisms that share recent common ancestry may have more in common than not.

1. \_\_\_\_\_
2. \_\_\_\_\_

Finally, list as many derived characters as possible that make *Homo sapiens* different from all other great apes. Be sure to restrict your list to truly BIOLOGICAL characters--not cultural ones. (This is where it gets really challenging, and sometimes there is simply not a clear line to draw, especially where cultural influences ("nurture") interact with a truly genetic and heritable ("nature") character.)

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_
4. \_\_\_\_\_
5. \_\_\_\_\_

As you can see, it is not a simple task to find biological characteristics that truly separate *Homo sapiens* from other species of great apes. In fact, we share more than 99% of our genes with our closest ape relatives, the Common Chimpanzees (*Pan troglodytes*) and Bonobos (*Pan paniscus*).

Take a look back at the several lists you have made, and note how synapomorphies identified at higher and higher resolutions help us to determine most recent common ancestry among the various taxa. Systematists use this method to construct and revise phylogenies for all living things.

### **3. Homologous vs. Analogous characters**

If the similarity between two characters in two separate taxa can be attributed to their presence in a common ancestor, then those two characters are said to be **homologous**. For example, the forelimb bones of all tetrapod (four-legged) vertebrates are homologous to one another, because they all evolved from the same bones in a common tetrapod ancestor. Although the bones may have evolved very different sizes, shapes, and functions, they all developed from the same embryonic sources and have evolved from the same ancestral tetrapod limb bones.

List five homologous characters you share with all other vertebrates that perform a similar function in you and all other vertebrates:

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_
4. \_\_\_\_\_
5. \_\_\_\_\_



Now list five homologous characters you share with other vertebrates that have evolved to serve a *different* function in you than they serve in some other vertebrates:

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_
4. \_\_\_\_\_
5. \_\_\_\_\_

Of the five characters you just listed, which are unique to *Homo sapiens*, and which are shared by at least some other vertebrates? What does this say about the recency of your common ancestry with those other vertebrates?

Not all physical similarities are homologous. In many cases, unrelated taxa have evolved superficially similar morphologies in response to similar natural selection pressure. For example, a shark and a dolphin both share a streamlined, **fusiform** shape well adapted for swift swimming. However, while the shark's body evolved from a fishlike ancestor with a somewhat fusiform shape, the dolphin's fishlike form is secondarily derived from that of a terrestrial, four-legged mammalian ancestor.

The superficial similarity of shark and dolphin is a result of **convergent evolution**. Specifically, what is meant by the term "convergent" evolution?

---

---

---

Characters that have evolved similar form and function from disparate ancestral sources are said to be **analogous**. Analogous characters are sometimes called **homoplastic characters** or **homoplasies** (from the Greek *homo*, meaning "alike" and *plas*, meaning "shape").

Don't let the similarity of the terms "homoplastic" and "homologous" confuse you! Look up their root derivations in your Dictionary of Word Roots and Combining Forms (Donald Borror). Write their exact, translated meanings here:

- homo* (Greek) = \_\_\_\_\_
- plas* (Greek) = \_\_\_\_\_
- analog* (Greek) = \_\_\_\_\_

List five characters you have that are analogous to characters with the same function but of different ancestral origin in any other species.

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_
4. \_\_\_\_\_
5. \_\_\_\_\_

## **B. Modern Systematics: The Cladistic System**

The **Cladistic System** of classification was first published by German zoologist **Willi Hennig** in 1950. The cladistic system is an objective and quantitative method of modern classification, and it is to its tenets that we—and the vast majority of modern systematists--adhere.

In the Cladistic System, organisms are classified exclusively on the basis of recency of descent from a common ancestor. Taxonomic/phylogenetic relationships are determined strictly on the basis of synapomorphies. The presence of many shared derived characters among taxa is considered

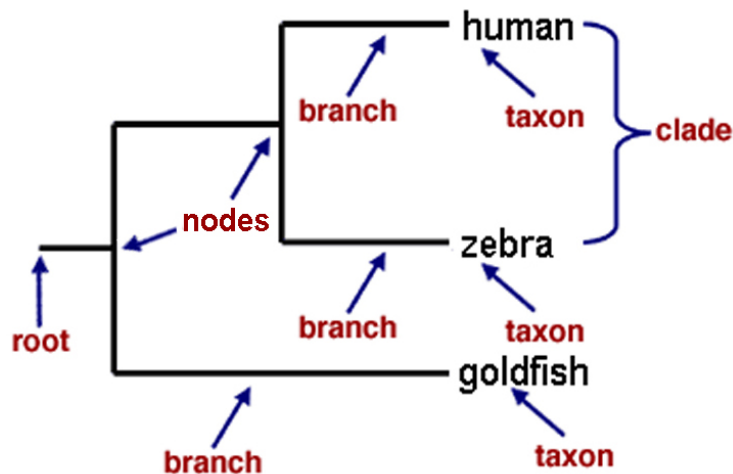
strong evidence of their common descent. Differences in the *degree* or *rate* of change among branches of organisms are irrelevant to their classification. For example, the cladist recognizes that birds--despite their plumage (modified scales homologous to reptile scales) and "warm-bloodedness"--share a more recent common ancestor with crocodilians than with other vertebrates. Birds and crocodilians (and their extinct relatives) are classified together in Archosauromorpha (literally "ancient lizard form"). In the old system (Classical Evolutionary Systematics), birds were placed in their own class ("Aves"), and crocodilians in a different class ("Reptilia") that did not reflect their common ancestry. Those two old names are now pretty much defunct, except when they are used as common names.

### C. Phylogenetic Trees

By considering synapomorphies and symplesiomorphies and by identifying homoplasies, the systematist attempts to construct phylogenies that reflect natural evolutionary relationships. Phylogenies can be represented as treelike diagrams showing how various extant taxa branched from common ancestors and from each other. Such a diagram is called a **phylogenetic tree** or **evolutionary tree** (Figure 2).

A phylogenetic tree shows putative evolutionary relationships. Taxa are represented as **branches** that bifurcate from **nodes**. A node represents the hypothetical common ancestor of the taxa on the branches above it on the tree, and it also represents the taxonomic unit (order, family, genus, etc.) that contains the taxa represented by the branches. Nodes are considered **hypothetical ancestors** (also known as **hypothetical taxonomic units**, or **HTUs**) because they cannot be directly observed. The base or **root** of the tree represents the hypothetical common ancestor of all the groups (clades) on the tree.

Phylogenetic trees may be represented in many different ways, some of which are shown in Figure 2. But the different diagrams all provide similar essential information.



**Figure 2. A phylogenetic tree.** The taxa included are "human," "zebra," and "goldfish." Each node represents the common ancestor of the taxa to the right of it on the tree. All taxa descended from a common ancestor comprise a clade. (Human and Zebra comprise a single clade; Human/Zebra/Goldfish comprise another, more inclusive clade that will have a higher taxonomic rank than the clade including only Human and Zebra.) Branches diverge from nodes, and represent a genetic unit descended from the ancestor at that node. The root represents the common ancestor of all taxa included on the tree.

The phylogenetic tree in Figure 2 shows only *recency of common descent*. It does not indicate which species might be (subjectively) described as "primitive" or "derived" (Those terms should be used to describe characters; they are meaningless when applied to an entire species.)

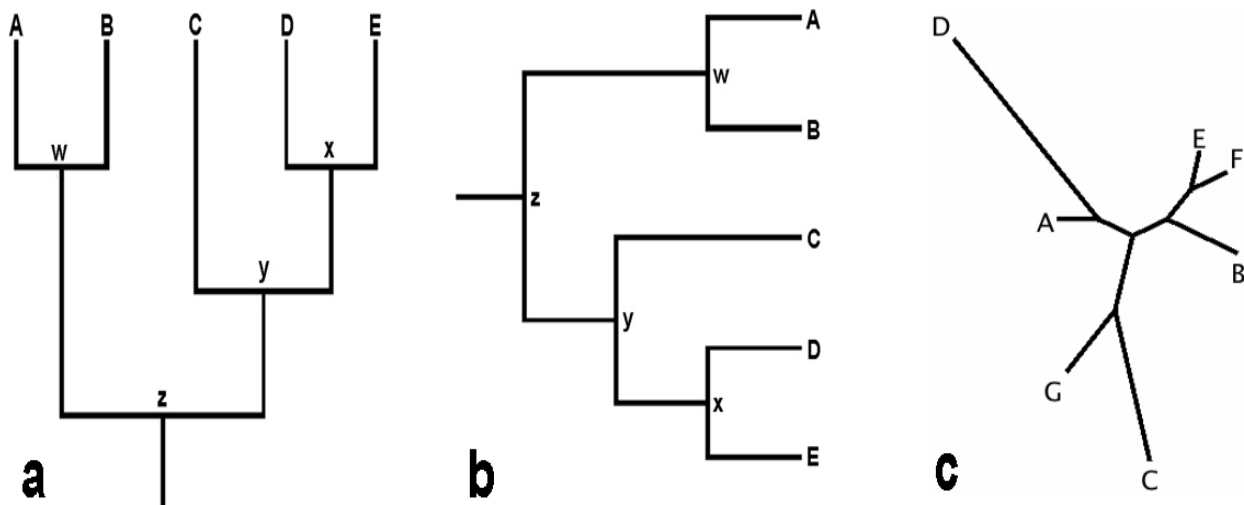
Note that two lineages branching from the same ancestor arose at the same geological time. Many people hold the misconception that *Homo sapiens* is the "most highly evolved" species, or even the most recently evolved. Neither is true. View Figure 3 below and remember the following rules.

**Rule #1: The branches at every node can be rotated.** The branches do not imply any sort of order; they indicate only recency of common descent. In Figure 4, the node at Ancestor F (Figure 4a) could be rotated so that the tree looked like the one shown in Figure 4b. The information given would be exactly the same. Any node on the tree can be rotated in a similar fashion.

**Rule #2: Two lineages branching from a single ancestral node are known as sister taxa.** Further specialization after a branch point is irrelevant. Therefore, it would be incorrect to say that humans evolved more recently than chimpanzees, or that humans should be placed in their own family simply because they seem so different from chimpanzees. *Taxonomic groupings are based on common ancestry only, not subjective perceptions of specialization.*

**Rule #3: There is no such thing as a "most highly evolved species".** All extant species are descended from successful ancestors, and are evolved to survive and reproduce in the context of their specific environment. Evolution is a process. It has neither a goal nor a subjective value system.

**Rule #4. No extant nor extinct taxon is considered ancestral to any other extant or extinct taxon.** Nodes represent hypothetical ancestors, not taxonomic units. When an ancestral lineage diverges to become two separate taxa, the ancestral lineage (hypothetical ancestor) is considered extinct, even if one of the descendant taxa is (or might be) virtually the same as that hypothetical ancestor. This should be remembered when one hears the oft-repeated, but incorrect statement "humans evolved from monkeys". They did not. Humans and monkeys share a common ancestor. That's not the same thing.



**Figure 3. Different graphic representations of phylogenies.** Whether a tree is upright, on its side, or drawn as an emerging spiral, the information is essentially the same. Nodes represent common ancestors of taxa (branches) above them on the tree, and endpoints represent taxa—whether extinct or extant—in the phylogeny.



- Figure 4a. Phylogeny of Primates.** The nodes from which branches emerge represent the hypothetical common ancestor of all taxa above that node on the tree. The endpoints of the branches represent the descendants of that ancestor. Some phylogenetic trees include both extinct and extant (still living) taxa. In modern systematics, extinct taxa (represented by fossils) are treated the same way as extant taxa, and are *not considered ancestral* to extant taxa.



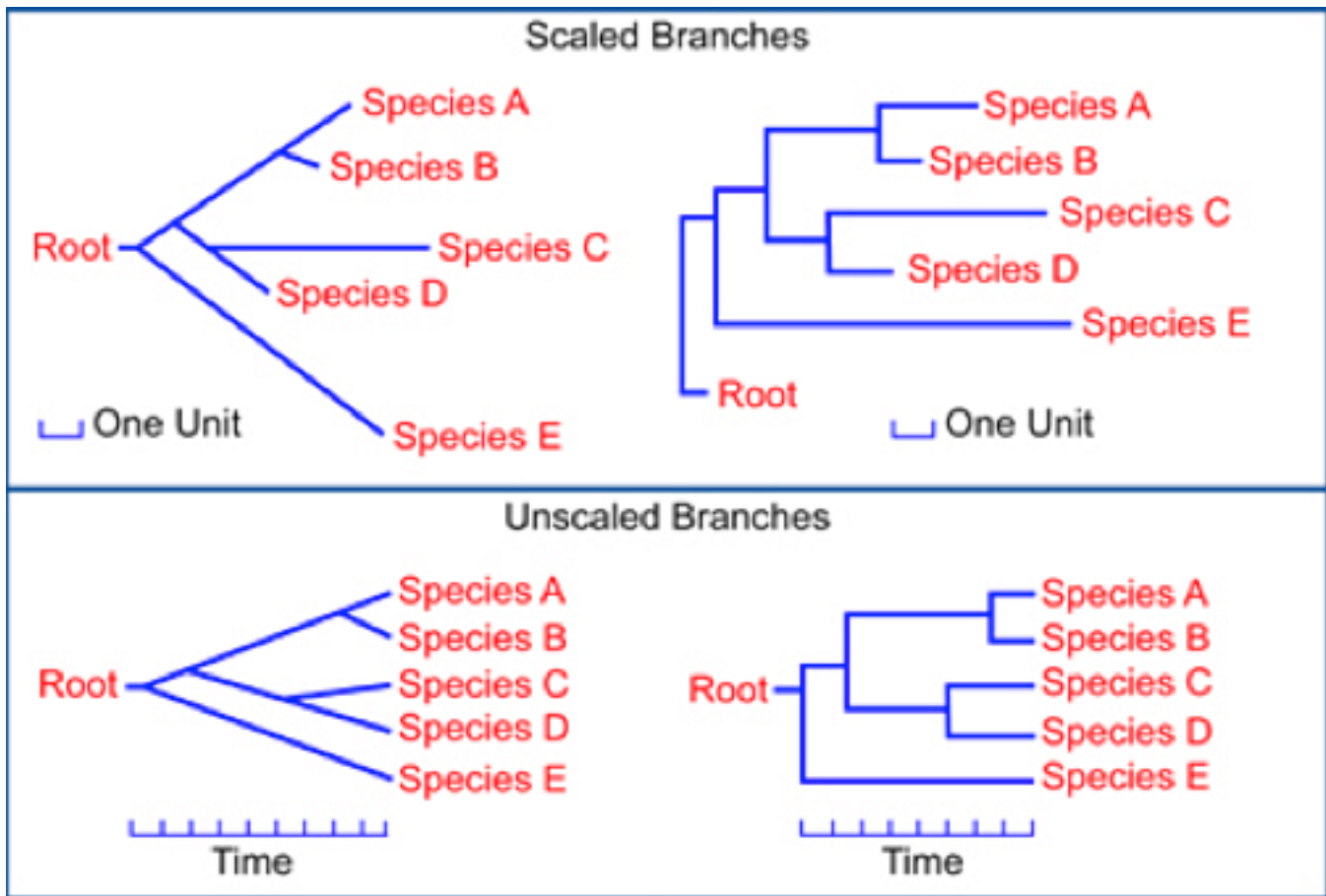
- Figure 4b. Phylogeny of Primates demonstrating a rotation of the node at Ancestor G, relative to the original drawing shown in Figure 3a.** Rotating the node in this manner does not change the phylogenetic information.

Figure 4 shows nine lineages of extant primates. Located beneath (to the left of) them on the tree are their hypothetical ancestors.

- The **Ancestral Primate** gave rise to all primates.
- Ancestor A** is gave rise to Tarsiers and Anthropoids, but not Lemurs and their kin.
- Ancestor E** is the most recent common ancestor of all Great Apes, but not Gibbons.
- Ancestor G** gave rise only to humans, chimpanzees and bonobos.

In trees such as those shown in Figure 4, the length of branches is not proportional to the degree of change from an ancestor or from related taxa. Such a tree is said to have **unscaled branches**. Time of divergence is shown by the time scale at the bottom of the diagram.

Modern systematics is done largely with molecular data, and it is now possible to quantify the degree of difference in DNA sequences. **Scaled branches**, such as those on the trees shown in Figure 5, indicate degree of divergence between taxa. (The unit is specified by the authors. For example, one unit might represent 1% of difference between two DNA or RNA sequences.)



**Figure 5. Scaled branches (top) can be used to indicate the degree of divergence between taxa on the tree, with respect to a particular character. Trees with unscaled branches indicate time of divergence with a time scale.**

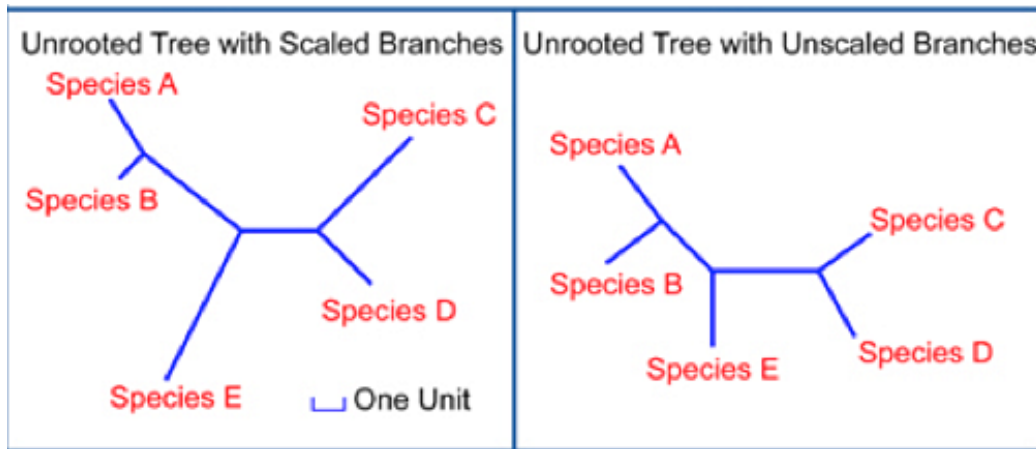
(Source: <http://www.ncbi.nlm.nih.gov/About/primer/phylo.html>)

The trees above all have a hypothetical ancestor their roots. Sometimes, however, the hypothetical ancestor is not known, and cannot be included on a phylogenetic tree.

### **Rooted and Unrooted Trees**

Rooting a tree is determining the location of the hypothetical common ancestor with respect to the other taxa on the tree. Because we have seen only rooted trees so far, this might seem like a no-brainer. But it's not always simple. Raw phylogenetic data often yield unrooted trees.

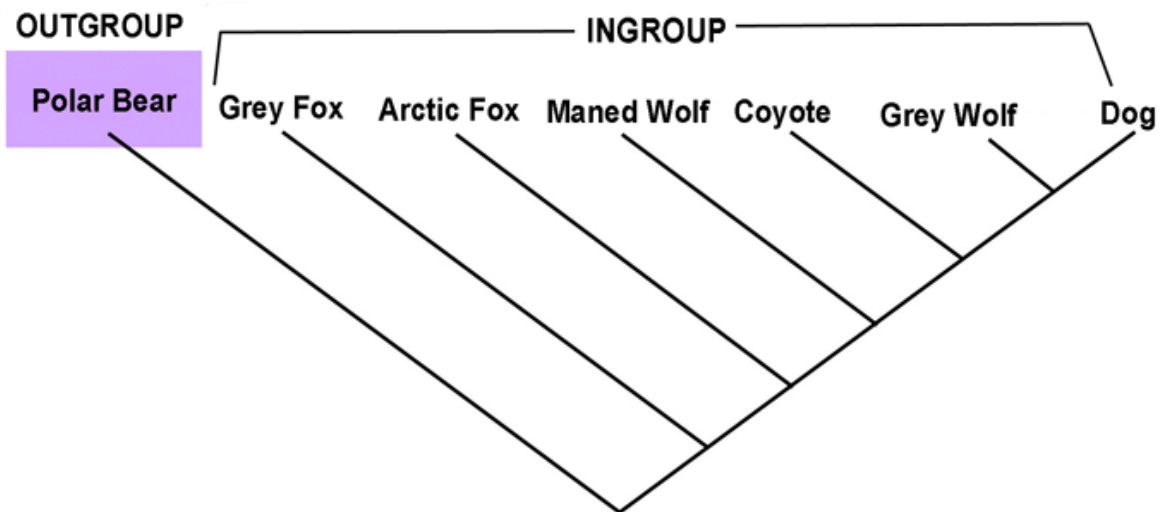
In an **unrooted tree** (Figure 6) there is no hypothetical ancestor, and no directionality to the tree. The tree shows only the putative evolutionary relationships of the taxa on the tree, without the evolutionary directionality implied by an ancestor.



**Figure 6.** Trees can be rooted or unrooted (Source: <http://www.ncbi.nlm.nih.gov/About/primer/phylo.html>)

In **rooted trees**, each node represents the most recent common ancestor of the taxa branching from it. Rooted trees are directional, with all taxa evolving or radiating from that single common ancestor at the root. In rooted trees, each ancestor is united to each node by a unique (evolutionary) path.

Clearly, unrooted trees are not as informative as rooted trees. In order to root a tree, one must consider a taxonomic unit that is closely related to, but phylogenetically *outside*, the group of taxa being studied. This closely related group is known as the **outgroup**. The outgroup is known to be more distantly related than the **ingroup** (the taxa of interest). It serves as a sort of “reference” group so that primitive characters common to taxa being studied can be identified: Characters common to both the outgroup and the taxa being classified are considered primitive to the entire assemblage. For example, if you were trying to determine symplesiomorphies of various species of Canidae (dog family), you might use a bear (Ursidae) as your outgroup (Figure 7).



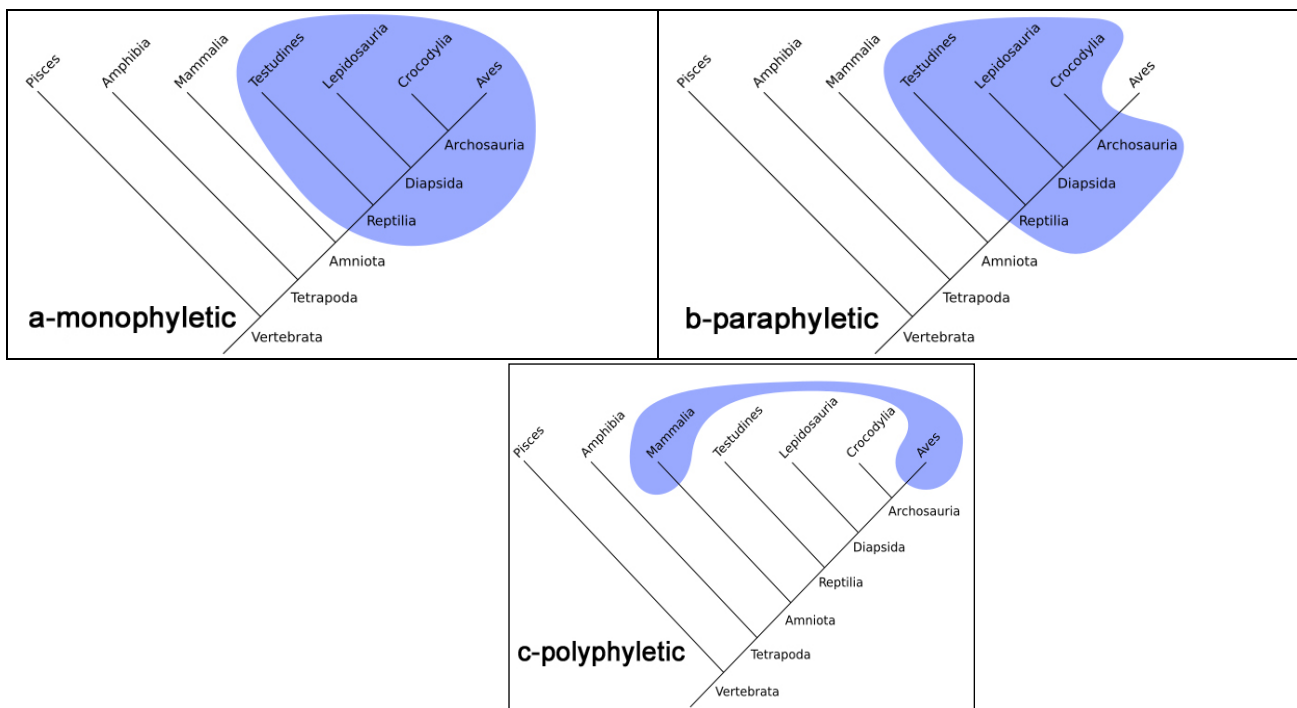
**Figure 7.** Hypothetical phylogeny of canid taxa, with member of Ursidae (Polar Bear) as outgroup. Characters common to the bear and the canids should be considered symplesiomorphies with respect to the canids of interest, and not used to construct less inclusive canid taxa.

## Monophyly, Polyphyly, Paraphyly

A phylogenetic tree is not constructed randomly. The systematist uses data on morphology, homology of nucleic acids, congruence of similar proteins, etc. to determine recency of common descent.

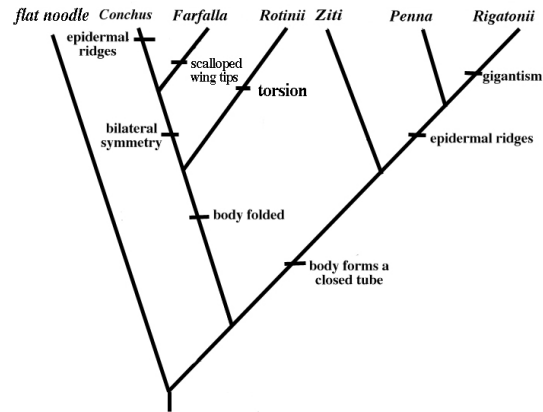
A **clade** is a group of species that includes an ancestral species and all of its descendants. Such a group is said to be **monophyletic**. The systematist uses cladistic techniques to construct monophyletic phylogenies that reflect true common ancestries. However, this is not always easy. When new data become available, it is sometimes discovered that a taxon under study is not monophyletic.

A **paraphyletic** taxon fails to include all descendants of a particular common ancestor. A **polyphyletic taxon** includes members that have descended from more than one different ancestor, but the common ancestor of those has not been included. These are illustrated in Figure 8.



**Figure 8.** Representative vertebrate taxa are grouped in monophyletic (a), paraphyletic (b) and polyphyletic (c) assemblages, shown by blue shading. Note that the paraphyletic tree (b) shows the traditional, evolutionary system for classifying Reptilia (turtles, crocodilians, snakes, and lizards), which does not reflect actual evolutionary relationships. Reptilia can be made monophyletic by including Aves (birds). The polyphyletic tree (c) illustrates what can happen when organisms are classified on the basis of superficial similarity, such as “warm bloodedness” or “four-chambered heart”. These characters most likely evolved independently in mammals and in birds.

Using some of the characteristics of the pasta you met earlier in this exercise, we have constructed a hypothetical phylogenetic tree showing their possible evolutionary relationships. (Figure 9) This may not be the only possible tree. The more data used, the more likely the tree will reflect actual evolutionary relationships. Notice that each character that sets a particular species apart from the others appears as a hashmark on the branch of the tree leading to that taxon.

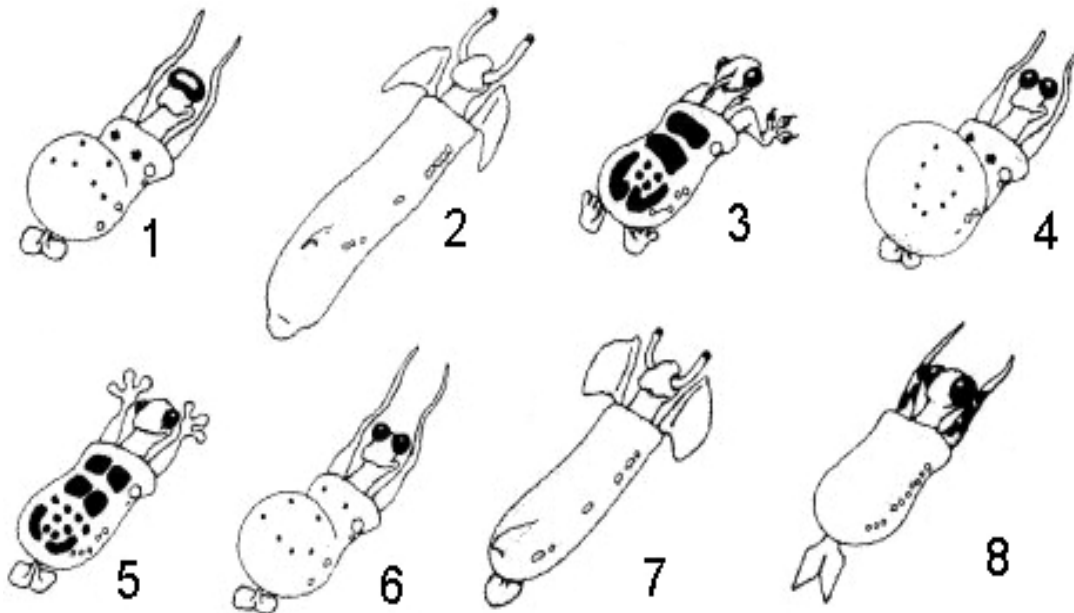


**Figure 9. A hypothetical phylogeny of the pasta of the United States.** The hash-marks along the tree indicate the appearance of synapomorphies found only in the taxa above that character on the tree. A flat noodle serves as the outgroup, indicating the primitive condition of the characters used to group the taxa.

### Exercise III. Constructing a Phylogenetic Tree

Classifications should be **natural**, meaning that they reflect evolutionary. We do not, for example, place slime molds and whales in the same family. Systematics, then, is a two-part endeavor. First, one must erect an hypothesis of evolutionary relationship among the organisms of interest. Second, one must devise a classification scheme that faithfully reflects the hypothesized relationship. We will use our Caminalcules to create a phylogenetic tree.

Examine the Caminalcules in Figure 10. These will be your Operational Taxonomic Units (OTUs)-a name we use to avoid assigning them to any particular taxonomic rank (such as species). Think of them as biological species, and refer to them by number.



**Figure 10. A variety of Caminalcules, arranged in no particular order.**



## Using Synapomorphies to Construct a Phylogeny

In the Cladistic System, the investigator groups OTUs together on the basis of synapomorphies. The presence or absence of a synapomorphy in two or more OTUs is inferred to be the result of inheritance (or lack thereof) from their common ancestor.

Results of a cladistic analysis are summarized in a phylogenetic tree called a **cladogram** (from the Greek *clad* meaning "branch"), an explicit hypothesis of evolutionary relationships. You already have seen an example of a cladogram, and know that monophyletic taxa are constructed on the basis of synapomorphies unique to each group. Now you'll get to do one yourself! Oo!

**Step One.** Select a series of characters that can be expressed as binary (i.e., two-state).

For example:

- Character a: "eyes present" (+) versus "eyes absent" (-)
- Character b: "body mantle present" (+) versus "body mantle absent" (-)
- Character c: "paired, anterior non-jointed appendages present" (+) versus "paired, anterior non-jointed appendages not present" (-)
- Character d: "anterior appendages flipperlike" (+) versus "anterior appendages not flipperlike" (-)
- Character e: "eyes stalked" (+) versus "eyes not stalked" (-)
- Character f: "body mantle posterior bulbous" (+) versus "body mantle posterior not bulbous" (-)
- Character g: "eyes fused into one" (+) versus "eyes separate" (-)
- Character h: "forelimbs with digits" (+) versus "forelimbs without digits" (-)

**Step Two.** Examine all your organisms and determine which character state it exhibits. Enter the data in a matrix like the one shown in Table 1.

Note that in this example, character *a* (presence or absence of eyes) and character *b* (presence or absence of a body mantle) is the same in all eight OTUs. Hence, this (primitive) character is not useful to us in finding differences between the OTUs.

Note also that only OTUs 2 and 7 share character *e* (stalked eyes), which is absent from all other OTUs. This suggests that OTUs 2 and 7 both inherited this character from a common ancestor. Likewise, OTUs 1, 4, and 6 share character *f* (bulbous mantle posterior), which is absent from all others. This supports the hypothesis of common ancestry among OTUs 1, 4, and 6. The same reasoning argues for common ancestry between OTUs 3 and 5 (character *h*), and so on.

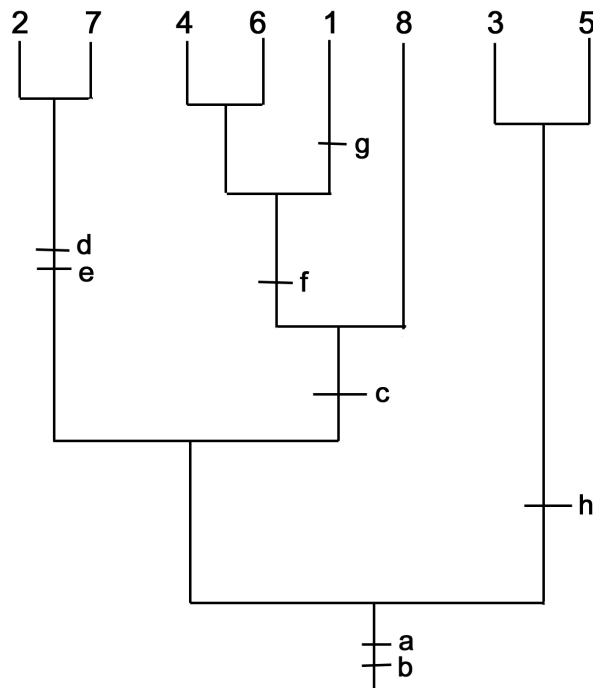
**Table 1. Character states of characters a - h in Caminalcules in Figure 10.**

character	1	2	3	4	5	6	7	8
<b>a</b>	+	+	+	+	+	+	+	+
<b>b</b>	+	+	+	+	+	+	+	+
<b>c</b>	+	+	-	+	-	+	+	+
<b>d</b>	-	+	-	-	-	-	+	-
<b>e</b>	-	+	-	-	-	-	+	-
<b>f</b>	+	-	-	+	-	+	-	-
<b>g</b>	+	-	-	-	-	-	-	-
<b>h</b>	-	-	+	-	+	-	-	-

A cladogram consistent with the distribution of these eight characters among the eight OTUs is shown in Figure 11.

This is not the only possible phylogeny consistent with the character distribution among the OTUs. In practice, there are often several--or even many--cladograms consistent with the data. In such cases, systematist generally applies a **parsimony** criterion for selecting the "best" cladogram. **The rule of parsimony states that when two or more competing hypotheses are equally consistent with the data, we provisionally accept the simplest hypothesis.** This is not to say that evolution is always parsimonious, only that our hypotheses should be.

In the case of competing cladograms, the rule of parsimony would require that we accept the simplest cladogram, the one with the fewest "steps" to each of the taxa on the tree. In our example, we could hypothesize that OTU 6 is actually more closely related to OTU 1 than to OTU 4. However, this would require that character g (fused eyes) had evolved once, and then secondarily lost in both OTUs 4 and 6. This is less parsimonious than stating fused eyes evolved only once, in OTU 1 only.



**Figure 11. A cladogram based on synapomorphies in Caminalcules 1 - 8.**

### **Cladistics and Linnaean Classification**

Given an hypothesis of evolutionary relationships, the second step in biosystematic endeavor is to erect a classification that faithfully reflects those relationships. Because the results of a cladistic analysis (i.e., the cladogram) are hierarchical, they can easily be incorporated into the Linnaean hierarchy, as shown in Figure 12. (Remember Linnaeus? If not, go back and review.)

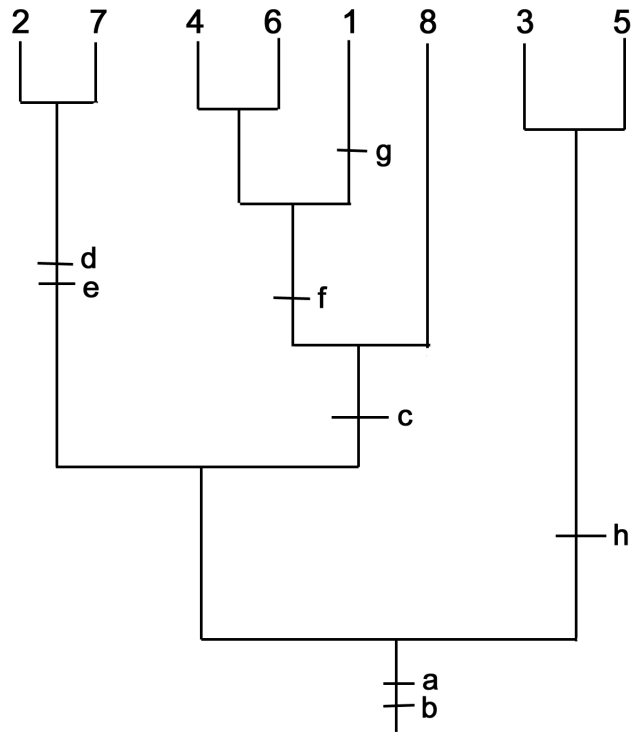
In cladistic analysis, all taxa must be **monophyletic**, meaning that they must include the common ancestor (almost always hypothetical) and all descendants of that common ancestor. Thus, in the cladogram above, OTUs 2 and 7 together with their common ancestor (at the branch point just below them) constitute a monophyletic genus, as do OTUs 1,4,6 and 8 and their common ancestor (at the branch point just above the appearance of character d).

A Family consisting of only OTUs 2 and 7 would not be monophyletic, because it does not include all the descendants of the common ancestor (at the branch point just below character d). Such a group would be considered **paraphyletic** (containing some, but not all, of a particular ancestor's descendants).

A Family consisting of OTUs 2 and 7 plus OTUs 3 and 5 would be considered **polyphyletic** (consisting of species derived from more than one most recent common ancestor). This is because such a taxon would be made up of groups descended from both the ancestor just below the appearance of character h, and the one just below the appearance of characters c and e.

**Order Caminalcula:**

- Family 1**
  - Genus 1**
    - Species 2
    - Species 7
  - Genus 2**
    - Species 1
    - Species 4
    - Species 6
    - Species 8
- Family 2**
  - Genus 3**
    - Species 3
    - Species 5



**Figure 12.** Incorporated results of a cladistic analysis showing Linnaean relationships among the OTUs.

**Exercise: Constructing a Cladogram Based on Synapomorphies**

Using the Caminalcules in the packet at your lab station, go through the steps of sample cladistic analysis we did for the Caminalcules in Figure 10. Use the Tables 3 and 4 to list shared, derived characters that help you group the OTUs into taxa that reflect their (hypothetical) evolutionary relationships. Finally, in the space provided, draw a cladogram of your Caminalcules, showing the appearance of each character, as in Figure 9. Is it rooted or unrooted? Be careful!

**Outgroup Analysis: Rooting the Tree**

As you may already realize, your phylogenetic tree of Caminalcules should be **unrooted** because you have no way to know which character states are derived. One is tempted to automatically assume that a more complex form is the more derived state. However, there is always the chance that a taxon has *lost* an ancestral character (or character state). (Can you think of any examples in our own species?)

To root your tree, you will need to consider the character states in an **outgroup**. The deck of Caminalcule cards your team will use has an outgroup on a light grey card. Your outgroup is one individual of a species of Caminalcules related to your ingroup (the rest of the Caminalcules on your cards). Use the outgroup to root your tree. Note that the rooted tree may be quite different from your unrooted one.

**Table 3. Character states in Caminalcule Packet.**

<b>character</b>	<b>state of character if (+)</b>	<b>state of character if (-)</b>
<b>a</b>		
<b>b</b>		
<b>c</b>		
<b>d</b>		
<b>e</b>		
<b>f</b>		
<b>g</b>		
<b>h</b>		

**Table 4. Synapomorphies in Caminalcule Packet.**

**OTUs**

<b>character</b>								
<b>a</b>								
<b>b</b>								
<b>c</b>								
<b>d</b>								
<b>e</b>								
<b>f</b>								
<b>g</b>								
<b>h</b>								

**A Cladogram of Caminalcules:**

**A Rooted Cladogram of Caminalcules:**

**You did it!**