

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 to be part of a particular taxon. From most inclusive to least inclusive, the major taxonomic **ranks** are as follows:

DOMAIN (e.g., Eukarya)
 KINGDOM (e. g., Animalia)
 PHYLUM (e. g., Chordata)
 CLASS (e. g., Mammalia)
 ORDER (e. g., Primates)
 FAMILY (e. g., Pongidae)
 GENUS (e. g., *Homo*)
 SPECIES (e. g., *Homo sapiens*)

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**. Within each of the major taxonomic ranks there may be larger and smaller taxa such as subkingdom, superphylum, subclass, subspecies, etc.

Every described, named organism is nested into a complete organizational hierarchy, from species through domain, as shown above for our own species, *Homo*

sapiens. Note that the scientific name of an organism (its genus and species) is *always* written with the genus capitalized and the **specific epithet** in lower case letters. 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 order to which all dogs belong (along with a host of other flesh-eating mammals with specialized cutting teeth called carnassials) is **Carnivora**.

The scientific name of a group of similar organisms has no more significance than any other convenient label used to describe a group of similar items. An example: the name of the taxon containing all domestic dogs is *Canis familiaris*. Taxonomic 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 you. 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 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 little pal is a flat-nosed frog with unwebbed toes!

2. The taxon's rank. Like the taxon's name, the taxon's rank has no true biological significance. It serves only to help the biologist locate the taxon within the hierarchy. An examples: The taxon "Eukarya" is assigned the rank of domain. The taxon "Mammalia" is assigned the taxonomic rank of class.

You may notice throughout this semester that a given taxon's rank may not always be the same in every source you read. For example, some publications may list "Zygomycota," "Ascomycota" and "Basidiomycota" as subphyla within Phylum Mycota, whereas others assign each of those three taxa the rank of phylum within Kingdom Fungi. Classifications shift as new data become available, but older publications are not changed to reflect the more recent classifications.

Confusing? We won't deny it. Just remember that as new data come to light, the ranks of familiar taxa may change with authors' attempts to create natural taxa.

The *relative* rank of a taxon within its larger and smaller groupings is more relevant than the rank itself, which is subject to change. 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 way, 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 as names change their rank.

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 specific 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. It's an ongoing quest--and one in which you might some day participate. Let's start with some simple exercises right now.

B. The Taxonomic Key: A Tool for Identification

Unfortunately, a biologist does not always receive materials neatly packaged with name and taxonomic information. In many cases, an investigator must identify an unknown specimen. A useful tool for this purpose 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**.

The 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 a container containing several "species" of pasta native to the United Aisles of Publix. The noodles have an evolutionary relationship to one another: they all are members of the same Order, Semolina, which evolved from a common ancestor resembling a soda cracker. A taxonomic key may or may not reflect these evolutionary relationships. It's simply a tool devised to allow identification of an individual organism to its proper taxonomic group. In this case, the key identifies each to genus and species.

Let's **key out** (this is a jargony verb commonly used to describe the process of identifying things with a taxonomic key) some pasta! Select one individual from your container, and use the taxonomic key below to identify its genus and species. Once you have done this, identify each different "species" of pasta in your container.

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

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, but many keys do so.

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 yourself. In another container at your station you will find several "species" of hardware. At this point, it's not important that you can tell their evolutionary relationships to one another. You are merely trying to create a tool that a field hardwareologist could use to identify them.

Create a dichotomous key to your hardware species, using the pasta key on the previous page as a guide for its construction. 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 up to you to decide.) Use proper *Systema naturae* rules in writing the scientific name of your hardware species: Genus capitalized, species lower case, and name italicized.

A Key to the Hardware of Southern Florida

1a. _____

1b. _____

2a. _____

2b. _____

3a. _____

3b. _____

4a. _____

4b. _____

5a. _____

5b. _____

6a. _____

6b. _____

Once you have finished your key, including all species of hardware, trade keys with the partners across the table from you. Using each other's keys, try to identify all hardware species. When done, check with your swap partners 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. Several different lines of evidence can be used to determine the degree of common ancestry between two taxa, including comparison of

1. gross morphology
2. biochemistry and physiology
3. chromosomes
4. cell ultrastructure
5. cellular metabolism and pathways
6. nucleic acid and protein sequences
7. geographical distribution (biogeography)
8. embryo development (ontogeny)

to name just a few. 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 and produce fertile, viable offspring under natural conditions.

1. Primitive and 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 arose on earth. We do know that extant species **evolved** by means of natural selection from previously existing **ancestral species**, and that this may involve **descent with modification** of traits (= **characters**) from one generation to the next. A character that shows little or no change from the same character in an ancestor is said to be **primitive**, whereas one that has changed in appearance and/or function relative to the ancestral form is said to be **derived**.

A primitive character is also known as a **plesiomorphy**, and a primitive character shared between two or more taxa is known as a **symplesiomorphy** (literally "shared primitive character"). A derived character is also known as an **apomorphy**, and a derived character shared between two or more taxa is known as a **synapomorphy** (literally "shared derived character").

All living things share 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)

To classify an unknown organism more precisely than just "a living thing," the systematist must consider characters that make that organism unique and different from members of other species. To this end, derived characters are the most informative characters to use. The next section explains why.

2. Symplesiomorphies and Synapomorphies

Because of the shared evolutionary history of all living things, each species shares certain very ancient (i.e., **primitive**, or **plesiomorphic**) characters with other species, whereas other more recently evolved (i.e., **derived**, or **apomorphic**) characters set them apart from other species. Synapomorphies that have been 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 recent their common ancestry.

For example, we humans share certain primitive characters with *all* members of Kingdom Animalia. List six primitive characters all humans share with all other animals, *but not with any other living things*:

1. _____
2. _____
3. _____
4. _____
5. _____
6. _____

Note: the characters you have listed above--if they are exhibited only by animals and by no other living organisms--are considered symplesiomorphies only with respect to Animalia. But if you are considering all living things, then the very same animal characteristics on your list are considered synapomorphies that set animals apart from all other living organisms. This means that any given character cannot be "primitive" or "derived" on its own. It can be described as "primitive" or "derived" only when compared to characters in other taxa.

With this in mind, now list three derived characters that set mammals (Mammalia, of which you are a member) 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.

You can see where this is going. Because you share those three characters with all mammals, they don't help you determine how closely related you are to any other mammal groups. To determine this, we must consider synapomorphies at an even higher resolution.

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 must be considered symplesiomorphies with respect to primates only. In other words, these three characters do 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 and more difficult to find synapomorphies linking particular taxa as they become smaller/less inclusive.)

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. _____
6. _____

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 DNA with our closest ape relatives, the Common Chimpanzees (*Pan troglodytes*) and Bonobos (*Pan paniscus*).

Take a look back at the several lists you have drawn, and note how synapomorphies identified at higher and higher resolutions help us to determine most recent common ancestry among the various taxa. You've done this with primates, but this is what biosystematists do with all taxa. This method is used to construct and revise complex phylogenies as new data become available.

3. Homologous and 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 be of 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 the same function in you as they do in 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**. In some sources, analogous characters are called "**homoplasies**" or "**homoplastic characters**." Before you allow the similarity of the two terms "homoplastic" and "homologous" confuse you, look up their root derivations in your Dictionary of Word Roots and Combining Forms. Write their exact meanings here:

homo (Greek) = _____
analog (Greek) = _____
plas (Greek) = _____

List five characters you have that are analogous to characters with the same function but of different ancestral origin in any other species. Discuss the evolutionary significance of each of these with your lab teammates.

1. _____
2. _____
3. _____
4. _____
5. _____

B. Schools of thought in Systematics

Not everyone agrees on a single correct way to classify organisms. Although the cladistic system is the most widely accepted school of thought in today's institutions of higher learning, there are still some holdouts. Take note of the tenets of each system.

1. Phenetic System (also known as Numerical Taxonomy)

Systematists who use the **Phenetic System** group organisms on the basis of phenotypic (physical) similarity. Pheneticists do not try to reconstruct evolutionary relationships, claiming that these relationships can never truly be known.

To devise a classification, the pheneticist chooses a number of phenotypic traits and determines whether they are present (+) or absent (-) in the organisms being studied. Groups having the most traits in common are said to be the most closely related. Such classification schemes usually identify symplesiomorphies, and do not use derived characters to construct phylogenies

A purely phenetic analysis often does not allow the investigator to distinguish homologous characters from homoplasies. This can result in the creation of taxa that fail to reflect evolutionary relationships (which the hard-core pheneticist would claim are irrelevant, anyway). Phenetics is rarely used in modern systematics.

2. Classical Evolutionary System

The **Evolutionary Classification System** is the one most commonly encountered by the student new to biology. Like the phenetic system, this classification groups organisms according to basic similarity, but unlike the phenetic system, it demands an evolutionary explanation for these similarities. Evolutionary taxonomists regard phenotypic specialization and degree of change *after divergence from a common ancestor* as important components of classification.

Traditionally, classical evolutionary taxonomists have considered a taxon worthy of separate status if its members show a high degree of specialization relative to those of a closely related taxon. Unfortunately, this can introduce subjectivity to the analysis.

Consider birds and reptiles. Molecular and ontogenetic data indicate that birds share a most recent common ancestor with crocodilians. However, because birds have feathers, are "warm blooded", and are superficially very different from crocodilians, classical evolutionary biologists placed them in Class Aves, and the crocodilians in Class Reptilia. This means that Class Reptilia does not include all the species that descended from the original ancestor that gave rise to lizards, snakes, crocodilians, and birds. Such an artificial taxon, which does not include all descendants of a single ancestor, is said to be **paraphyletic**

Similarly, *Homo sapiens* was once assigned to its own family (Hominidae), although there is no objective reason to taxonomically separate it from the great ape family (Pongidae). A Family Pongidae that does not include *Homo sapiens* is paraphyletic, as it does not include all descendants of the common ancestor of great apes. This subjective separation of humans because they seem somehow “special” obscures true evolutionary relationships.

3. The Cladistic System

The **Cladistic System** was first published by German zoologist **Willi Hennig** in 1950. Systematists who use this system classify organisms 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 common descent.

The cladist holds that differences in the rate of evolutionary 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 most recent common ancestor with crocodilians. The modern cladistic classification places birds and crocodilians (and their extinct relatives) together in a new taxon, Archosauromorpha (literally “ancient lizard form”).

Although some systematists feel that the cladistic system's weakness is its failure to consider unequal rates of evolution and specialization in a taxon after it has speciated from its relatives, the cladist would argue that these are cladism's greatest strengths. The cladistic system is the most objective and quantitative of modern classification systems, and it is to its tenets that we—and the vast majority of modern systematists--adhere. The other two schools of thought are of historical interest only.

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 off" from common ancestors and from each other. Such a diagram is called a **phylogenetic tree** or **evolutionary tree**.

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 (Figure 1)

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

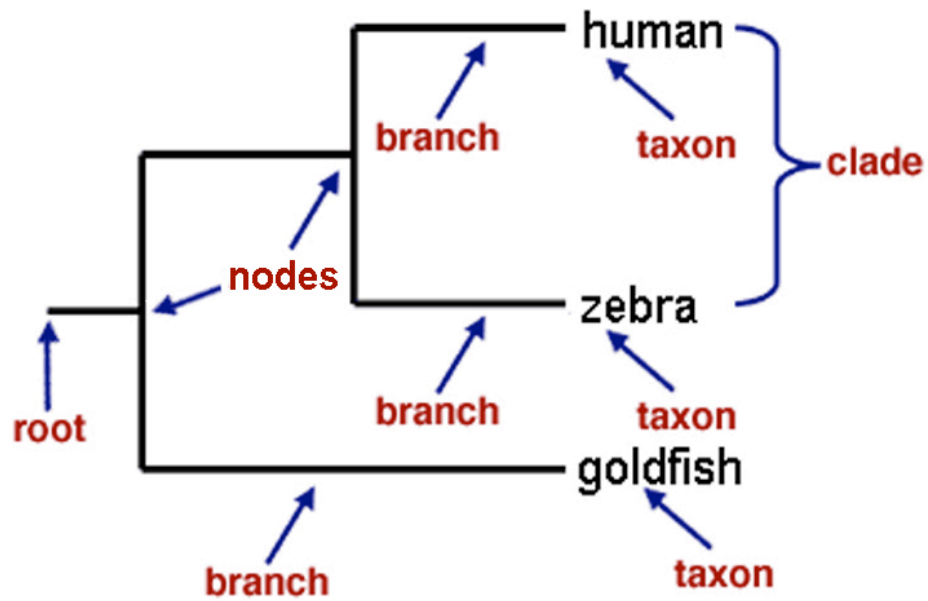


Figure 1. A phylogenetic tree. The taxa included are “human,” “zebra,” and “goldfish.” Each node represents the common ancestor of the taxa to the right of it on the tree. All taxa descended from a common ancestor comprise a clade. (Human and Zebra comprise a single clade; Human/Zebra/Goldfish comprise 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.

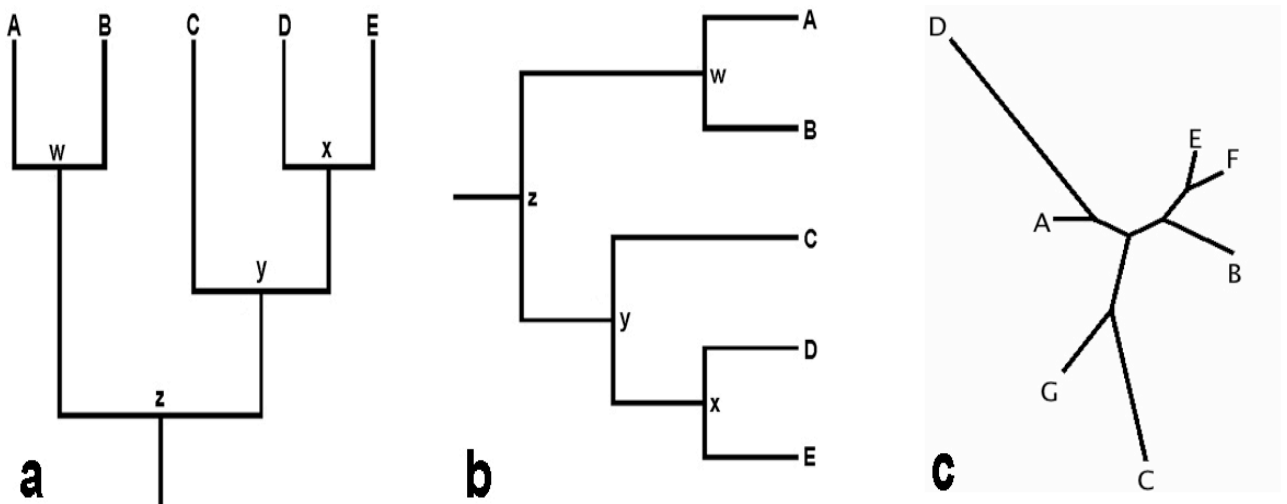


Figure 2. Different graphic representations of phylogenies. Whether a tree is upright, on its side, or drawn as an emerging spiral, the information is essentially the same. Nodes represent common ancestors of taxa (branches) above them on the tree, and endpoints represent taxa—whether extinct or extant—in the phylogeny.

A Sample Phylogeny: Primates

Figure 3 shows nine extant lineages of 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 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 only rise to humans, chimpanzees and bonobos.

Note that this phylogenetic tree shows only *recency of common descent*. It does not indicate which species might be (subjectively) described as "primitive" or "derived" (Those terms are meaningless when applied to an entire species.)

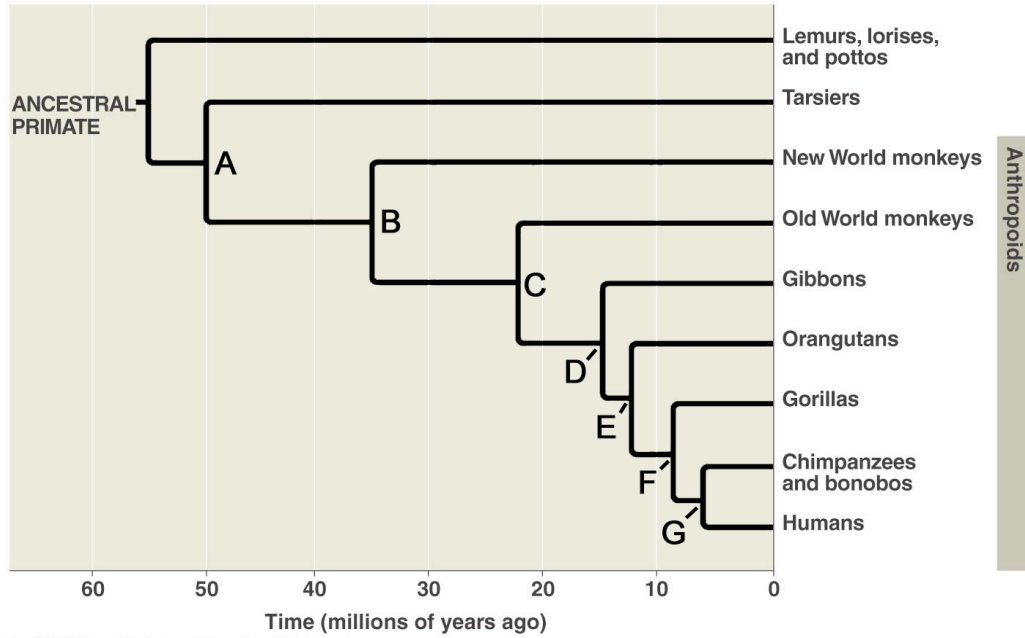
Note also that two lineages branching from the same ancestor arose at the same geological time. Many people have the misconception that *Homo sapiens* is the "most highly evolved" species, or even the most recently evolved. Neither is true. Always remember the following rules.

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. For example, the node at Ancestor F could be rotated so that the tree looked like the one shown in Figure 3b, and 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 to the systematist using cladistic analysis (which we do). Therefore, it would be incorrect to say that humans evolved more recently than chimpanzees, or that Humans should be placed in their own family ("Hominidae") simply because they seem so different from chimpanzees. Taxonomic groupings are based on common ancestry only, not subjective perceptions of specialization.

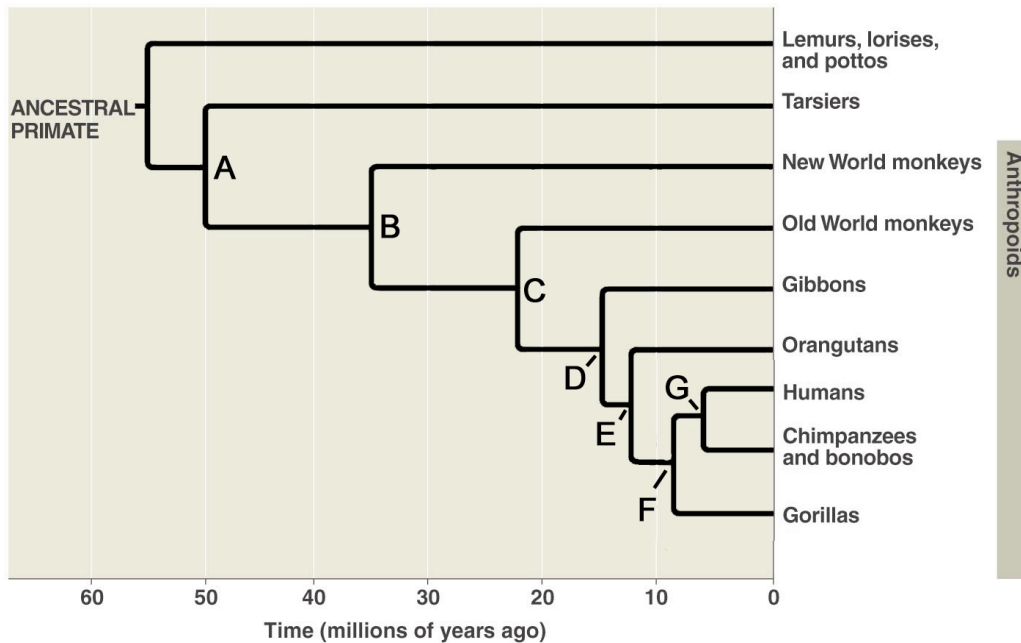
Rule #3: There is no such thing as a "most highly evolved species". All extant species are descended from successful ancestors, and are evolved to survive and reproduce in the context of their specific environment. Evolution is a process. It has neither a goal nor a subjective value system.

Rule #4. No extant or 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 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.



Copyright © 2008 Pearson Education, Inc., publishing as Pearson Benjamin Cummings.

Figure 3a. 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.



Copyright © 2008 Pearson Education, Inc., publishing as Pearson Benjamin Cummings.

Figure 3b. 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.

Rooted and Unrooted Trees

The trees in Figure 3a and 3b show the evolutionary relationships of Primates as they branched into their respective taxa from a hypothetical ancestor at the root of the tree. Sometimes, however, the hypothetical ancestor is not known, and not included on the tree.

In a **rooted tree** (Figure 4), 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.

In an **unrooted tree** (Figure 4) 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.

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**.

In a cladistic phylogenetic analysis, the systematist uses the state of homologous characters to group taxonomic units together on the basis of synapomorphy. The character state of the appropriate outgroup is likely to represent the primitive, ancestral condition of that character.

In some phylogenetic analyses, it is relevant to indicate the degree of change in a particular branch (taxon) as compared to others. Longer branches indicate more change, and shorter branches, less change. A tree with branches calibrated to show degree of change is called a **scaled tree**. Sometimes, but not always, a time scale is included with a tree to indicate how long ago divergence from an ancestor occurred. (Figure 4).

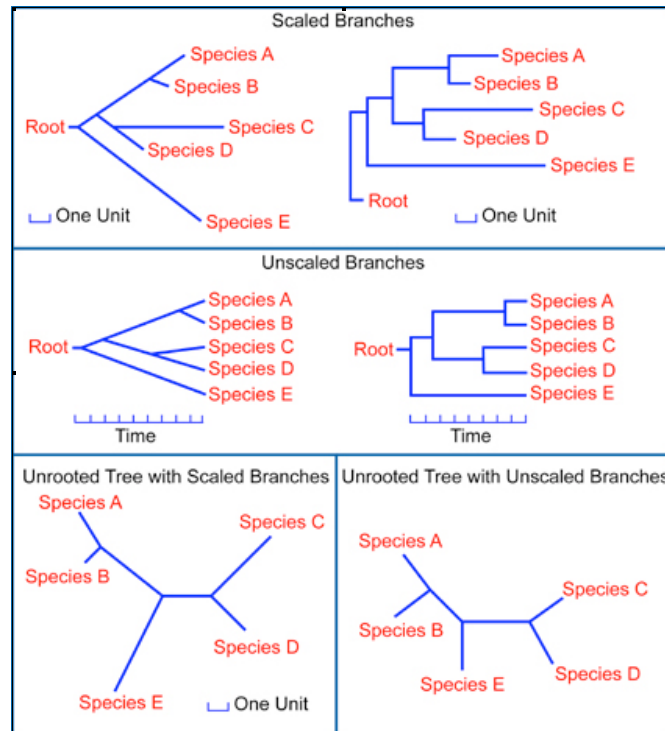


Figure 4. Trees can be rooted or unrooted, scaled or unscaled, or combinations of both. (Source: <http://www.ncbi.nlm.nih.gov/About/primer/phylo.html>)

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 5.

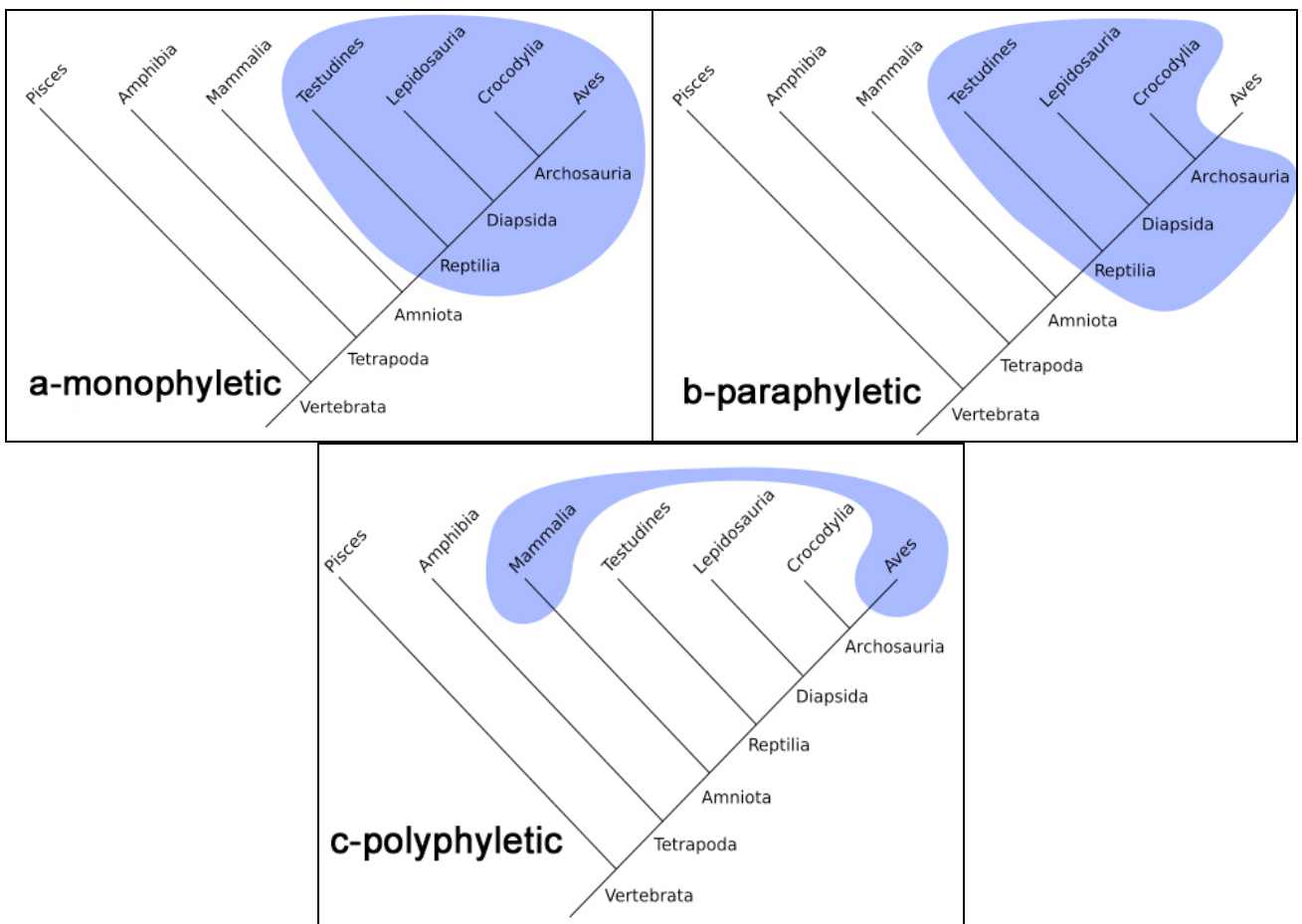


Figure 5. Representative vertebrate taxa are grouped in monophyletic (a), paraphyletic (b) and polyphyletic (c) assemblages, shown by blue shading. Note that the paraphyletic tree (b) shows the traditional, evolutionary system for classifying Reptilia (turtles, crocodilians, snakes, and lizards), which 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 6) This may not be the only possible tree, and the more different data sets used to construct a tree that show congruency between them, the more likely it is that the tree reflects actual evolutionary relationships.

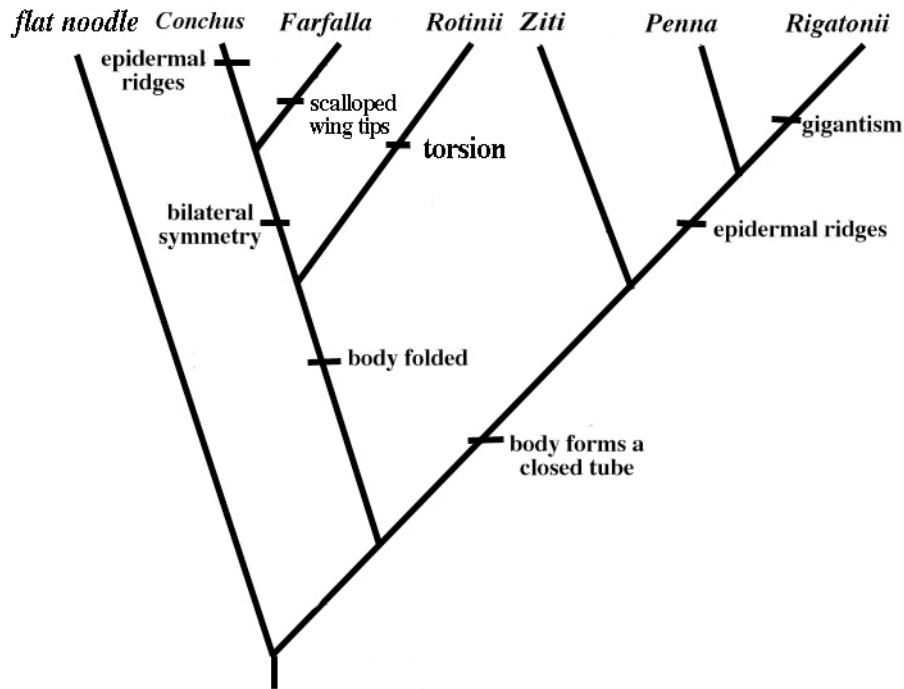


Figure 6. A hypothetical phylogeny of the pasta of the United States. The hashmarks along the tree indicate the appearance of synapomorphies found only in the taxa above that character on the tree. A flat noodle serves as the outgroup, indicating the primitive condition of the characters used to group the taxa.

Exercise IV. Constructing Phylogenetic Trees

Most biologists agree that our classifications should be "natural" so that to the extent possible they reflect evolutionary relationships. We do not, for example, place slime molds and whales in the same family. Biosystematics, then, is a two-part endeavor. First, one must erect an hypothesis of evolutionary relationship among the organisms under study. Second, one must devise a classificatory scheme that faithfully reflects the hypothesized relationship. We will use a series of imaginary animals to introduce you to two rather different methods that attempt to do this.

The hypothetical animals used in this exercise are called **Caminalcules** and were created and "evolved" by **J. H. Camin**, Professor of Biology at the University of Kansas in the 1960's. They have served as test material for a number of experiments concerning systematics, its 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.

Phenetics: Cluster Analysis and Phenograms

In a phenetic analysis, organisms are clustered on the basis of overall similarity. In this portion of the lab, you will perform a cluster analysis on Caminalcules (Figure 7) to understand the assumptions of this approach, its strengths and its weaknesses.

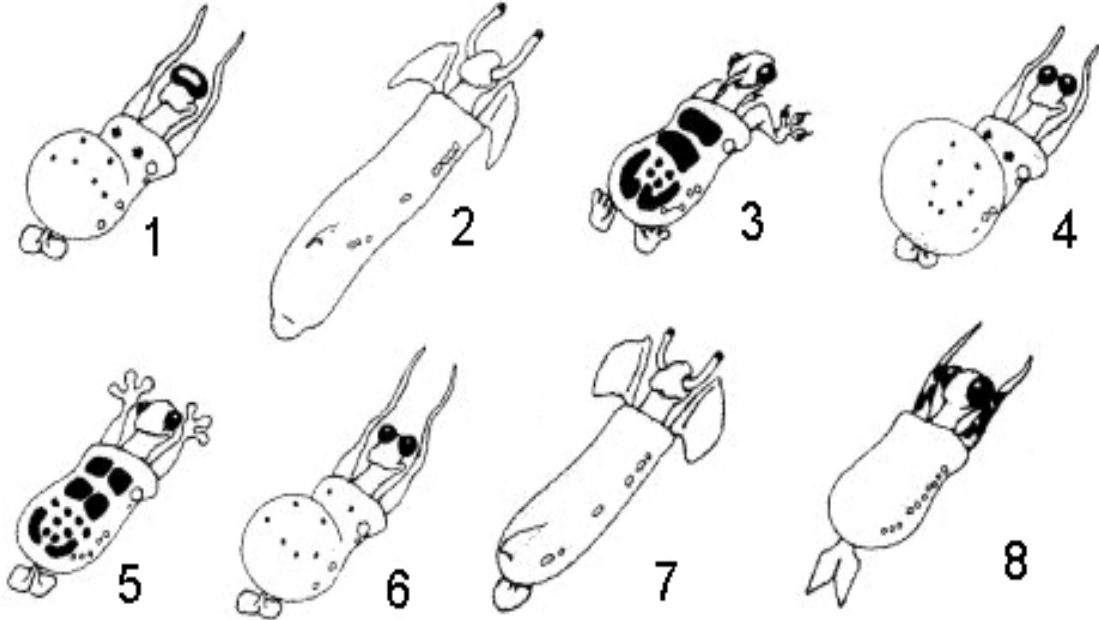


Figure 7. A variety of Caminalcules, arranged in no particular order.

A Sample Procedure

Examine the Caminalcules in Figure 7. 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.

Make a subjective judgment about the overall similarity between all pair-wise combinations of the eight OTUs, using a scale of 1.0 (maximum similarity) to 0 (complete dissimilarity). These rankings are then cast into a similarity matrix. An example of such a matrix (for the eight OTUs in Figure 7) is shown in Table 1.

Table 1. An example of similarity rankings among the eight Caminalcules pictured in Figure ST-4. The rankings have been subjectively assigned.

	1	2	3	4	5	6	7	8
1	-							
2	0.1	-						
3	0.2	0.1	-					
4	0.7	0.3	0.4	-				
5	0.5	0.2	0.8	0.3	-			
6	0.8	0.2	0.4	0.7	0.4	-		
7	0.1	0.9*	0.2	0.3	0.2	0.3	-	
8	0.5	0.3	0.6	0.4	0.6	0.4	0.4	-

Step 1. Find the pair of OTUs that have the highest similarity ranking. (In this example, it happens to be OTUs 2 and 7, with a similarity ranking of 0.9 shown in **boldface blue with an asterisk**).

Step 2. Combine OTUs 2 and 7, and treat them as a single composite unit from this point on. Construct a new matrix (this time it will be 7 x 7), as shown in Table 2.

Table 2. Reduced matrix with similarity values recomputed for all OTUs with composite OTU 2/7.

	2/7	1	3	4	5	6	8
2/7	-						
1	0.1	-					
3	0.15	0.2	-				
4	0.3	0.7	0.4	-			
5	0.2	0.5	0.8*	0.3	-		
6	0.25	0.8*	0.4	0.7	0.4	-	
8	0.35	0.5	0.6	0.4	0.6	0.4	-

Step 3. Recalculate the similarity values for each OTU with the new composite 2/7 OTU. To do so, simply compute the average similarity of each OTU with 2 and with 7 (listed in Table 1).

In our example, the similarity of OTU 4 with OTU 2 is 0.3, and the similarity of OTU 4 with OTU 7 is 0.3, then the similarity of OTU 4 with OTU 2/7 will be the average of those two similarities: $[(0.3 + 0.3)/2]$, or 0.3. The recalculated similarities for each OTU with composite OTU 2/7 are shown in **boldface** in Table ST-2. The new highest similarity value(s) is/are marked in **boldface blue with an asterisk***.

Step 4. In the new, reduced matrix with recomputed similarity values, find the *next* pair of OTUs with the highest similarity value. In this case, OTUs 1 and 6 and OTUs 3 and 5 are tied, similarity values of 0.8. For simplicity, choose one pairing at random and recalculate the similarity indices, and then do the next pairing, as shown in Table 3(a).

Step 5. Continue to construct reduced matrices, each time recalculating the similarity indices between your new composite OTU with all remaining OTUs, as shown in Table ST-3(b - e). The last step will result in a 2 x 2 matrix with a single, final similarity value. (In this example, composite OTY 4/1/6/8/3/5 has a 0.3 similarity to composite OTU 2/7).

Step 6. Your OTUs can now be clustered graphically in a branching diagram called a **phenogram**. The result of our sample cluster analysis is shown in Figure 8. Note that the most similar OTUs from the first table (OTUs 2 and 7) have been paired at a branch point reflecting their similarity index (0.9). The next most similar OTU pairings (1/6 and 3/5) have been clustered at their respective levels (0.8) in the same fashion, and each successive reduced matrix yields the appropriate branch level, shown on the similarity scale at the left side of the diagram.

This phenogram is designed to not only indicate which of the Caminalcules are most physically similar, but also *to what degree* they are phenotypically similar.

Table 3. Reduced matrices with similarity values recomputed for all OTUs with (a) new composite OTU 1/6, (b) new composite OTU 3/5, (c) new composite OTU 4/1/6, and (d) new composite OTU 8/3/5 and (e) 4/1/6/8/3/5. Again, recalculated similarity indices are shown in boldface, and the new highest similarity index in each matrix is marked with an asterisk.

a:

	1/6	2/7	3	4	5	8
1/6	-					
2/7	0.18	-				
3	0.3	0.15	-			
4	0.7	0.3	0.4	-		
5	0.45	0.2	0.8*	0.3	-	
8	0.45	0.35	0.6	0.4	0.6	-

b:

	3/5	1/6	2/7	4	8
3/5	-				
1/6	0.38	-			
2/7	0.35	0.18	-		
4	0.35	0.7*	0.3	-	
8	0.6	0.45	0.35	0.4	-

c:

	4/1/6	3/5	2/7	8
4/1/6	-			
3/5	0.37	-		
2/7	0.24	0.35	-	
8	0.43	0.6*	0.35	-

d:

	8/3/5	4/1/6	2/7
8/3/5	-		
4/1/6	0.4*	-	
2/7	0.35	0.24	-

e:

	4/1/6/8/3/5	2/7
4/1/6/8/3/5	-	
2/7	0.3*	-

Note again that the similarity values assigned in the example were subjective. Different relative similarities might have been assigned by a different observer. A phenogram provides a visual representation of similarity relationships of OTUs, *and nothing more*. Note that clusters can be rotated around their branches without changing the implied similarity relationships. For example, if the diagram were rotated at the 0.3 branch point so that OTUs 2 and 7 were on the right, and OTUs 1, 3, 4, 5, 6 and 8 on the left, the information in the diagram would not be changed. The same would be true of rotation at any branch point.

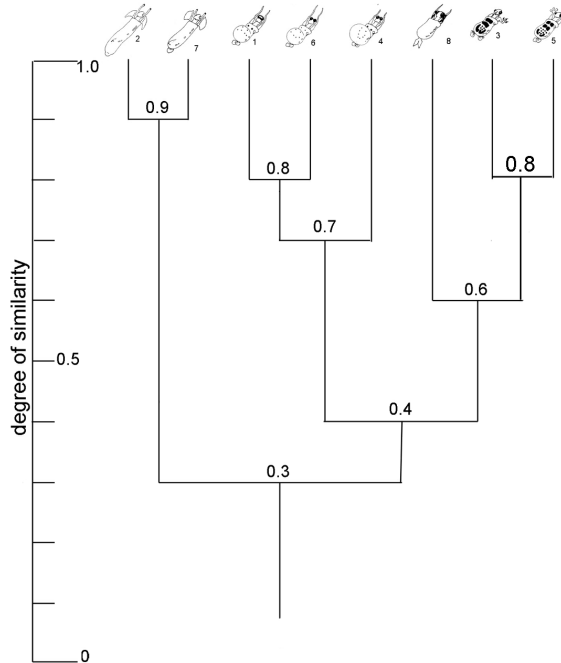


Figure 8. Phenogram of Caminalcules clustered in the sample exercise.

Given the computational tedium of working through this simple 8 OTU example, you can see why real numerical taxonomic studies, which often involve large numbers of OTUs, require the use of computers.

Exercise: Constructing a Phenogram Using Cluster Analysis

You will now do your own cluster analysis and construct a phenogram using a somewhat different group of Caminalcules from your sample OTUs. A small deck of laminated cards, each showing a unique Caminalcule, is available at your lab station. You will use these as your OTUs.

A series of tables is provided for you to do your reduction calculations, and space on the following page is provided so that you may draw a phenogram for your Caminalcules. Work in pairs for this exercise. Once you have finished the exercise, discuss the questions following the exercise.

Questions

1. In the early years of numerical taxonomy, advocates of the approach claimed that the methodology would bring objectivity to systematics by the application of highly quantitative techniques, specific algorithms for clustering, and numerical assessment of similarity. What do you make of these claims of objectivity?
2. An underlying assumption of numerical taxonomy is that overall similarity is a reliable indicator of evolutionary relationships (i.e., recency of common ancestry) among organisms. Is this a reasonable assumption?
3. If numerical taxonomy is such a flawed method for inferring evolutionary relationships, why are we torturing you with all this? (Actually, this is a serious question: Do you believe that cluster analysis might be of use in other areas of biology? If so, in what fields, and how?)

Matrices for your cluster analysis:

	-							
		-						
			-					
				-				
					-			
						-		
							-	
								-

Cladistics: Synapomorphies and Cladograms

Cladistic methods provide a more objective and suitable way of inferring evolutionary relationships. Cladistics is sometimes known as **quantitative phyletics**. Rather than grouping OTUs on the basis of overall similarity, as in numerical taxonomy, the investigator using this method groups OTUs together on the basis of shared, derived characters (**synapomorphies**)--characters whose presence or absence in two or more OTUs is inferred to be the result of inheritance from their common ancestor.

Results of a cladistic analysis are usually 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 in Figures 3a and 3b, and know that monophyletic taxa are constructed on the basis of synapomorphies unique to each group. You'll now have a chance to do it yourself!

A Sample Procedure

We will examine the eight Caminalcules in Figure 7, this time in an attempt to infer the evolutionary relationships among them.

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 4.

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 if we are trying to determine 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 these three OTUs.. The same reasoning argues for common ancestry among OTUs 1, 4, and 6 (character f), and for OTUs 3 and 5 (character h), and so on.

Table 4. Character states of characters a - h in Caminalcules in Figure 4.

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

A cladogram consistent with the distribution of these eight characters among the eight OTUs is shown in Figure 9.

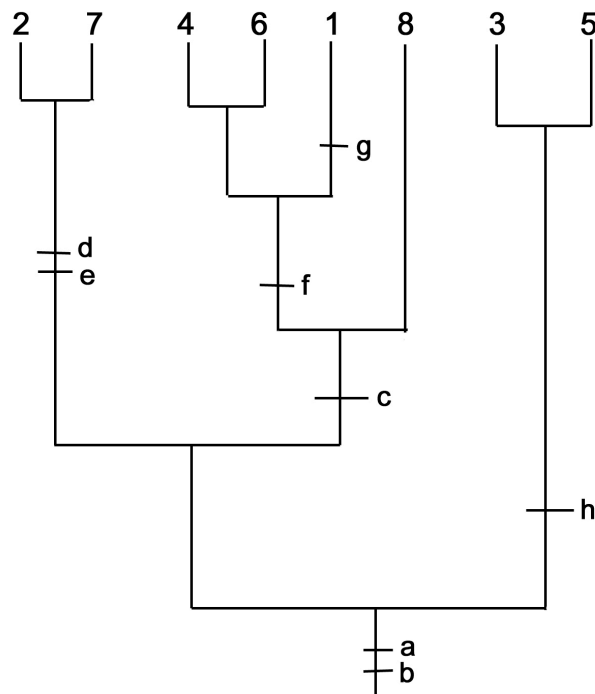


Figure 9. A cladogram based on synapomorphies in Caminalcules 1 - 8.

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 that can be constructed, all of which are 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 been 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.

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 10.

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

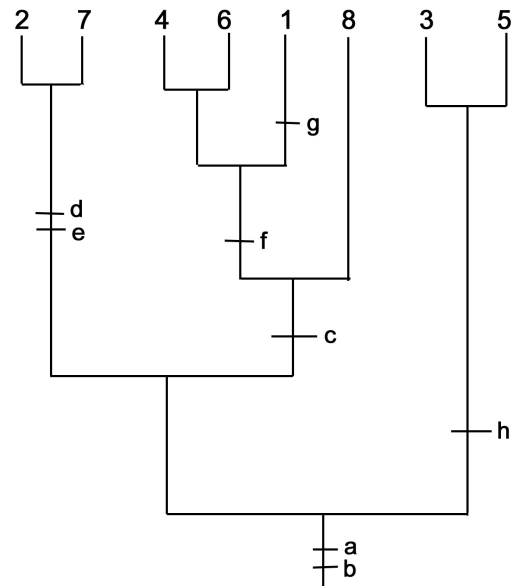


Figure 10. Incorporated results of a cladistic analysis showing Linnaean relationships among the OTUs.

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 7. Use the spaces and matrix provided to choose shared, derived characters that help you group the OTUs into taxa that reflect their hypothetical evolutionary relationships. Finally, in the space on the next page, draw a cladogram of your Caminalcules, showing the appearance of each character, as in Figure 10. Is it rooted or unrooted? Be careful!

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

OTUs

character								
a								
b								
c								
d								
e								
f								
g								
h								

Outgroup Analysis: Rooting the Tree

As you may already realize, the phylogenetic tree you created in the last exercise should be an **unrooted** tree 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 possibility that a species (or other taxon) has *lost* a character (or character state) that was present in the ancestor. (Can you think of any examples?)

To root your tree, you will need to consider the character states in an outgroup. But how? Fortunately, we have some for you right here in the lab.

Once you have completed your initial analysis and your cladogram has been examined and approved by your laboratory instructor, select a member of your team to go to the front desk and randomly select a card from the envelop labeled "outgroups."

Each card has an illustration representing one individual of a species of Caminalcules related to yours, but not in the same taxa. Use this outgroup to root your tree. Note that the rooted tree may be quite different from your unrooted one.

A rooted cladogram of Caminalcules:

Questions

1. The results of phenetic and cladistic analyses are inherently hierarchical, as is the branching sequence of the evolution of organisms. So, too, is the Linnaean classificatory system of ever more inclusive taxonomic categories from species, Genus, Family, Order, Class, Phylum, Kingdom and Domain. Can you name some other hierarchical classification systems (not necessarily biological)?
2. Taxonomists were erecting classifications of organisms long before Darwin convinced biologists of the reality of evolution. Some of these taxonomists believed in the fixity of species and in special creation. Nevertheless, in some respects, these pre-Darwinian classifications are rather similar to those produced later by evolutionary taxonomists. Why do you think this is so?