

How to Read a Phylogenetic Tree

Deborah A. McLennan

Published online: 29 September 2010
© Springer Science+Business Media, LLC 2010

Abstract It has been over 50 years since Willi Hennig proposed a new method for determining genealogical relationships among species, which he called phylogenetic systematics. Many people, however, still approach the method warily, worried that they will have to grapple with an overwhelming number of new terms and concepts. In fact, reading and understanding phylogenetic trees is really not difficult at all. You only need to learn three new words, autapomorphy, synapomorphy, and plesiomorphy. All of the other concepts (e.g., ancestors, monophyletic groups, paraphyletic groups) are familiar ones that were already part of Darwinian evolution before Hennig arrived on the scene.

Keywords Phylogenetic tree

Dan Brooks and I teach a biodiversity course (EEB 265) to second year students at the University of Toronto. The entire course is structured around a phylogenetic framework. We begin with the big, albeit simplified, tree of the Metazoa, then work our way from sponges to snakes, focusing on the characters that bind groups together and the characters that make each group unique. If we are doing our job correctly, our students should be able to answer the following questions—what is this animal (how do you know)? What does it do? What makes it special? What aspects of its biology make it vulnerable to anthropogenic intervention? Since all of the students had already taken a lab in first year biology covering the fundamentals of phylogenetics, we assumed that we

wouldn't need to review phylogenetic methodology in our biodiversity course. It didn't take long for us to realize that our assumption was naïve; by the time many of the students had arrived in EEB 265, they had already hit the delete button next to “phylogenetics” in their brain. It is always humbling to (re)discover that not everyone shares your views about the things in life that are interesting and important!

Back to the drawing board. One of the major problems with teaching a course about metazoan diversity is that you simply don't have enough time to cover all of the groups. The last thing we wanted to do was to sacrifice biology-based lectures for a discussion about theory. So, the challenge was simple: design a lecture that would, in 50 minutes, teach students how to understand what a phylogenetic tree was telling them. It wasn't our intention to teach students how to make trees, just how to read them. This paper is based on that lecture.

The word “phylogeny” is a combination of two Greek words, *phyle* (tribe—in particular, the largest political subdivision in the ancient Athenian state [www.yourdictionary.com; www.etymonline.com]; another word we get from this is “phylum”) and *geneia* (origin [www.etymonline.com]; another word we get from this is “gene”). It was coined by the developmental biologist Ernst Haeckel in 1866 and then championed by Darwin in his famous work, *On the Origin of Species* (beginning with the 5th edition in 1869). Both biologists tied the idea of “phylogeny”—the origin of groups—to evolution. Phylogenetic trees are thus simply diagrams that depict the origin and evolution of groups of organisms.

Although you might not know it, we are all familiar with the idea of phylogenetic trees. People have been making such trees for decades, substituting the word, “family” for “phylogenetic” (Fig. 1). Just as individual people in a family over generations are connected by bonds of “blood” (the process of reproduction that produces offspring),

D. A. McLennan (✉)
Department of Ecology and Evolutionary Biology,
University of Toronto,
25 Willcocks St.,
Toronto, ON M5S 3B2, Canada
e-mail: deborah.mclennan@utoronto.ca

Family trees represent historical “ties of blood”
(= genes = reproduction)

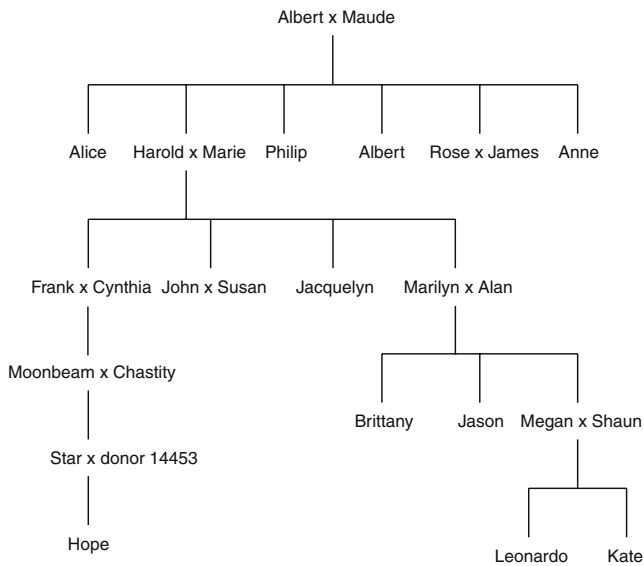


Fig. 1 Family tree for an interesting group of people. In phylogenetic terms, family trees (genealogies of people) = phylogenetic trees (genealogies of species)

individual species are connected by evolutionary ties (biological processes like natural selection and geological processes such as continental drift or a river changing course that produce species). In this sense, speciation (the production of new species) = reproduction (the production of new individuals). In other words, we are all, from members of the same family to members of the same species, connected by genes.

Family trees tend to be drawn as if they were hanging upside down, like a cluster of grapes. Phylogenetic trees are depicted somewhat differently. Imagine that you are holding the family tree for the big cats shown in Fig. 2a. Now, flip it sideways (rotate 90° counterclockwise) and you have the image shown in Fig. 2b. Rotate this image yet another 90° counterclockwise, smooth it out, and you have the image shown in Fig. 2c (this tree shape was the one used by Darwin in *On the Origin of Species*). The important thing to remember is that all three depictions are saying exactly the same thing about the relationships among species of big cats. How you choose to draw your phylogenetic trees depends, in part, on personal preference—some people find it easier to read 2b, others prefer 2c.

Phylogenetic trees are reconstructed by a method called “phylogenetic systematics” (Fig. 3). This method clusters groups of organisms together based upon shared, unique characters called *synapomorphies*. For example, you share the presence of a backbone with cats, but not with butterflies. The presence of a backbone thus allows us to hypothesize that human beings are more closely related to cats than they are to

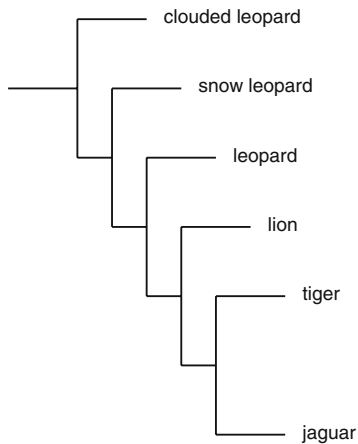
butterflies (Fig. 4a); cats and people both have a backbone, butterflies are spineless¹. Not all characters are synapomorphies. Some traits, called *plesiomorphies*, are shared by all the members of a group. Returning to our tree, we see that cats, people, and butterflies all have DNA (Fig. 4b). The presence of DNA allows us to hypothesize that these three species are all part of the same group, but it does not tell us anything about how those species are related to one another. Think of it this way: my last name tells me that I am part of the McLennan clan. If I meet someone called Jessie McLennan, I know we are related somehow, but I haven’t any idea whether she is a long lost cousin or someone from a more distant branch of the family tree. The final term you need to know is *autapomorphy*—traits that are only found in one member of the group. For example, butterflies can be distinguished from cats and people because they have an exoskeleton made out of chitin (a tough, waterproof derivative of glucose). Autapomorphies help us identify a particular species in a group but, like plesiomorphies, they tell us nothing about relationships within the group. Overall these three types of characters can be likened to the story of Goldilocks: plesiomorphies are too hot (too widespread), autapomorphies are too cold (too restricted), and synapomorphies are just right (for determining phylogenetic relationships).

Enough of characters for the moment; back to the trees themselves. Why do the branches on a tree have names (e.g., lion, tiger, etc.), while the lines joining different branches together do not (Fig. 5)? This is because these lines represent *ancestors*. An ancestor is a species that has undergone a speciation event to produce descendant species. The ancestor usually “disappears” in the process of speciation. Does this mean that the ancestor goes extinct?

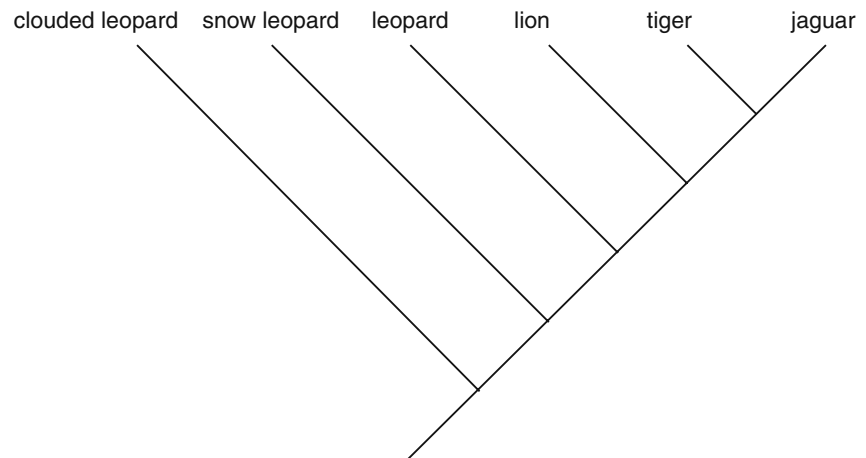
In order to answer this, we must do some time traveling carrying a digital device that records everything we see (Fig. 6). Imagine you travel back 10,000,000 years, then stop, intrigued by an interesting species of lizard with red spots all over its back (species A). After a while, you decide to move forward in time five million years or so then stop again. You search around and discover two new lizard species, one with blue spots on its back (species B), and the other with red stripes (species C), but species A is nowhere to be seen. Did it go extinct? You look back over your digital recording of those five million years and discover that species A split into two groups, which became different in some ways from one another through time. In evolutionary terms, species A is an ancestor (ancestor 1) and species B and C are its descendants. Fast forward to today (with more digital material to watch) and you find three

¹ Photograph of monarch butterfly courtesy of Daniel Brooks; image of the Simpsons from Simpsons wallpaper (www.simpsonstrivia.com.ar/simpsons-wallpaper.htm); Kitten is Taala (which means “wind” in the language of the Nuxalk Nation people, Bella Coola, BC)

So many ways to draw a family (phylogenetic) tree



(b) another way to draw the same tree



(c) yet another way to draw the same tree

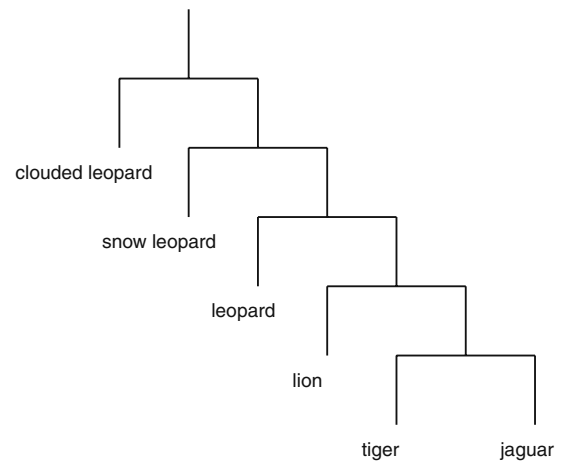
Fig. 2 a–c So many ways to draw a family/phylogenetic tree for the genus *Panthera*

species of lizard: your old friend the blue spotted lizard (species B) and two new lizards (descendants of species C, the red striped lizard), one with blue stripes (species D) and the other with a solid black back (species E). Today, then, there are only three species of lizard alive. You no longer see either of the ancestors (the red spotted and red striped lizards), but we still show them on the phylogenetic tree.

The answer to our original question “did the ancestor go extinct?” is thus No! In many cases, the ancestor is subdivided and the biological (genetic) information encompassed within the ancestor is passed on to the descendant species. Over time, the descendants change and become different in some ways from each other and from the ancestor, while retaining some things in common (for example, all of our lizard species have a backbone). This is evolution.

So what really counts as extinction? *Extinction is the loss of biological information—the physical loss of a species.* For

(a) usual way family trees are drawn for people (in this case you see a family tree for big cat species)



example, consider a simplified phylogenetic tree of the dinosaurs (Fig. 7). All of the groups on dotted branches are extinct—none of the species in those groups exist on this planet anymore (Jurassic Park notwithstanding), which means that all of the information that was *unique* to each of those groups has been lost. The only group that managed to avoid extinction was Aves (or birds)—avian species are the last remaining dinosaurs.

OK, let’s take what we have learned about ancestors and clustering groups based on shared, unique characters (synapomorphies) and use that to decipher the information contained within a phylogenetic tree. Here is a tree depicting the relationships among living members of the Amniota, a large group of vertebrates that includes most of the animals with which you are familiar (Fig. 8). You already know that the names of species, or groups of species, are written across the tips of the branches on the

Phylogenetic trees are reconstructed via a method called phylogenetic systematics

* this method clusters taxa on the basis of **shared, unique characters (synapomorphies)**

for example, you share the presence of a backbone with cats, but not with butterflies



butterfly (*Danaus plexippus*)



cat (*Felis catus*)



human beings (*Homo sapiens*)

Fig. 3 The basis of phylogenetic systematics

tree. The next thing you need to know is that characters are depicted at their point of origin on a phylogenetic tree. So, on this tree you can see that (1) the amniotic egg originated in ancestor 1 and was passed on to all of its descendants (mammals, ancestor 2, turtles, ancestor 3, ancestor 4, crocodiles, birds, ancestor 5, tuataras, and lizards plus snakes). In evolutionary terms, the amniotic egg is a unique trait that is shared only by ancestor 1 and all of its descendants; (2) a special type of skin protein (β keratin) originated in ancestor 2 and was passed on to all of its descendants (turtles, ancestor 3, ancestor 4, crocodiles, birds, ancestor 5, tuataras and lizards plus snakes). β keratin is a unique trait shared by the group called “Reptilia”; and (3) a breakable tail originated in ancestor 5 and was passed on to all of its descendants (tuataras, lizards plus snakes). A breakable tail is a unique trait shared by members of the group tuataras + lizards + snakes.

In fact, every organism is a complex mosaic of thousands of traits. If you don’t believe this, sit down and list all of the traits that make you, you. In addition to the obvious things like eye color and hair color, don’t forget the fact that you have RNA, DNA, individual cells, an anterior and posterior end, a skull, jaws, bone, arms and legs, come from an amniotic egg, have three bones in your inner ear, were suckled on milk produced in mammary glands, have an opposable thumb, and no tail. In other words, when you look at a phylogenetic tree, you will see that all of the branches have at least one, and more likely many, characters on them (the slash marks on Fig. 9a). Because of this, it is often difficult to actually label all of the traits on a tree because it’s visually distracting. A shorthand method has been developed to deal with this problem: draw the tree showing the relationships

among the groups (Fig. 9b) and list the synapomorphies for each branch elsewhere in a table. On the other hand, if you are interested in one or more particular traits, you can highlight them on the phylogenetic tree without showing all the other characters. For example, if you wanted to discuss the evolution of mammals, you could show the amniote tree and highlight just the synapomorphies for the mammals (e.g., three middle ear bones: Fig. 9c). Remember, this is just shorthand!

There is one last thing about characters that is important to understand: characters are not static things. They evolve through time. In other words, a “synapomorphy” may not “look the same” in all species that have it. So, for example, consider the stapes, one of the three bones in your middle ear that are responsible for transferring sound waves from the eardrum to the membrane of the inner ear. This small bone has a long, complicated, and fascinating evolutionary history. To understand that history, we must travel back many of hundreds of millions of years to the origin of the Deuterostomes, a large group that includes the Echinodermata (starfish and their relatives), Hemichordata (worm-like, marine creatures), and Chordata (amphioxus + tunicates + Craniata [organisms with skulls]). The ancestor of this large group had numerous slits in its pharynx (called visceral arches) that were involved with filter feeding. Time passed and cartilaginous rods providing support for the arches appeared, were subdivided and modified. The upper section of the second visceral arch rod is the focus of our tale (Fig. 10). As we move forward still further in time, this character undergoes various structural and positional modifications; in essence, it becomes larger, more robust, and involved in supporting the jaws (at which point it is called the hyomandibula), changes from cartilage

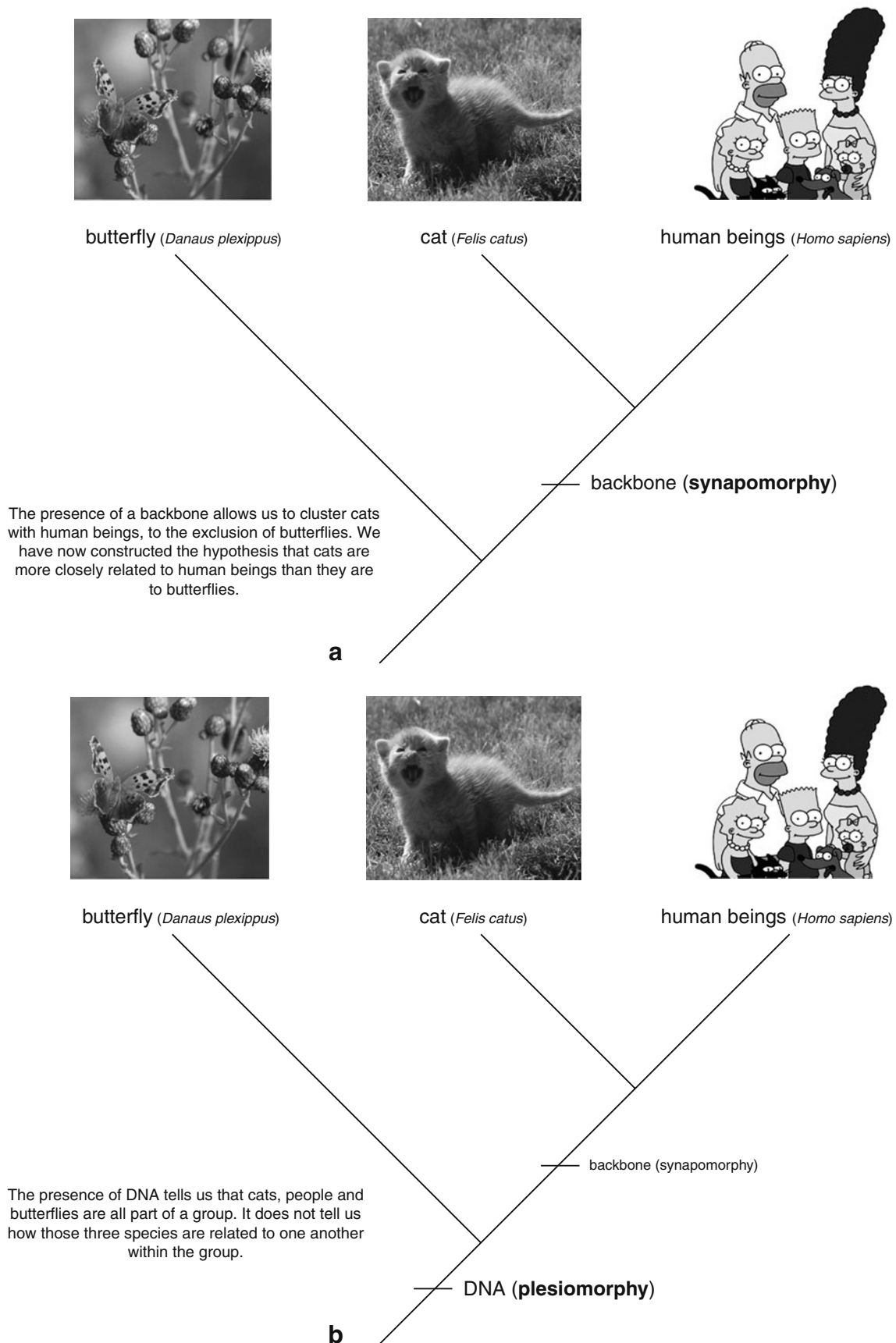


Fig. 4 Identifying types of characters on a phylogenetic tree. **a** a synapomorphy; **b** a plesiomorphy; **c** an autapomorphy

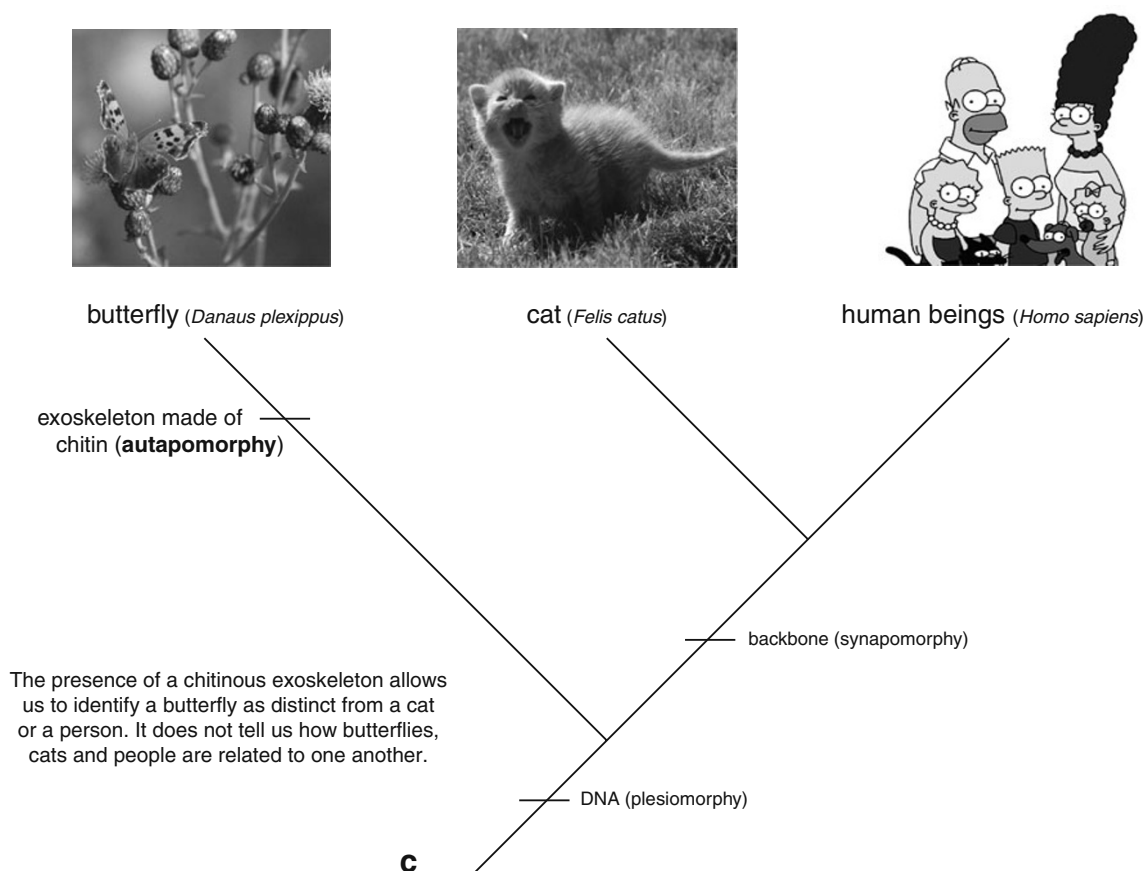
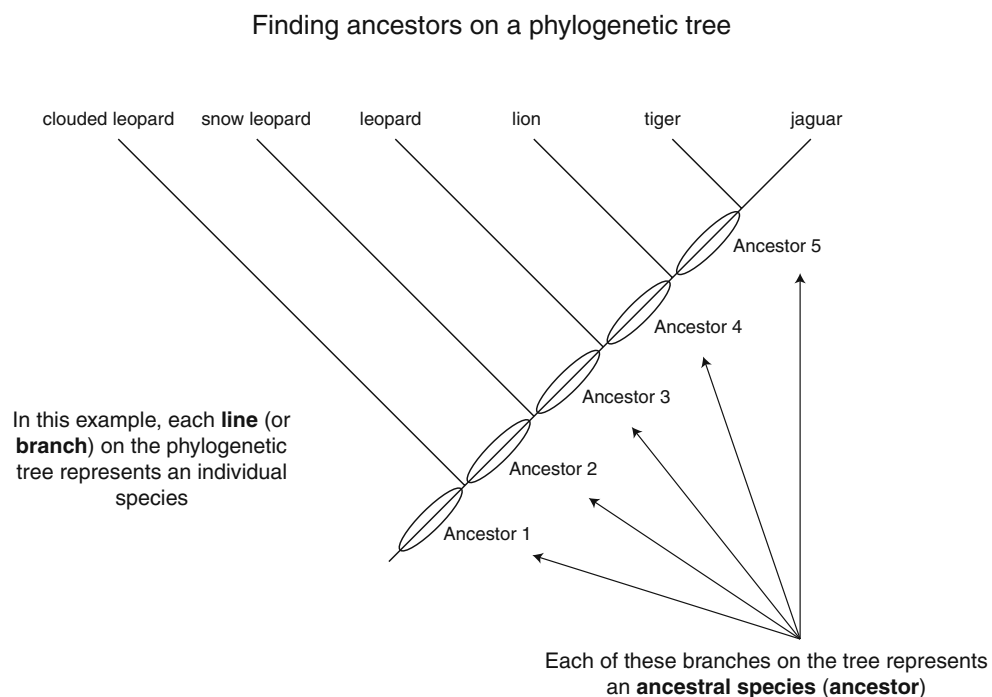


Fig. 4 (continued)

Fig. 5 Finding ancestors on a phylogenetic tree



Are the ancestors extinct?

Time travel, speciation and ancestors

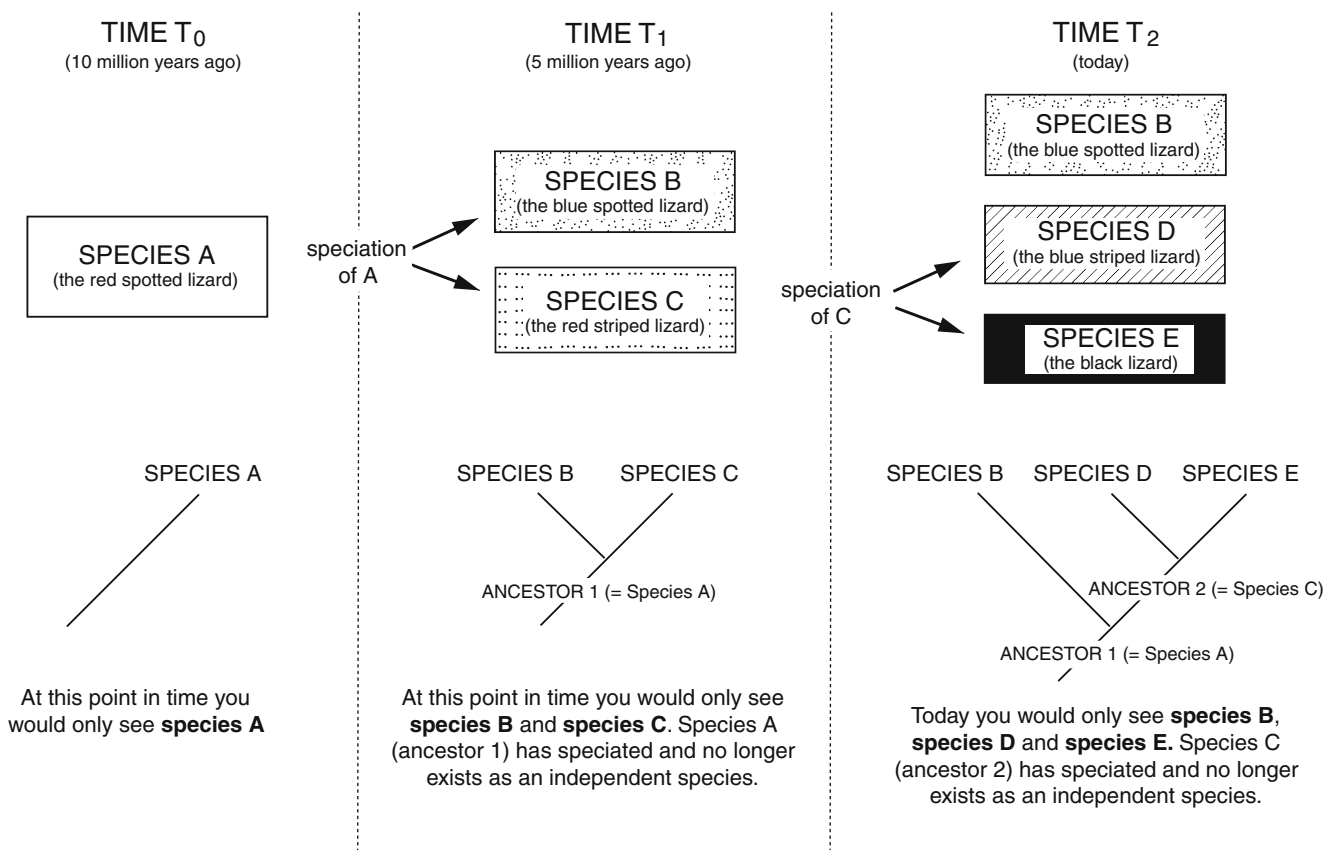


Fig. 6 Traveling back in time to discover ancestors

to bone, then begins a gradual reduction in size, disengages from the jaw/cheek area, and moves into the middle ear (at which point it is called the stapes). Overall then, the upper portion of the 2nd visceral arch—hyomandibula—stapes is the same structure that has had both its shape and function modified over hundreds of millions of years. So although the presence of a “cartilaginous rod in the 2nd visceral arch found in the throat region” may be a synapomorphy for the Craniata, you won’t find that exact structure in any four-footed animals. Instead, what you will find is the modification of that cartilaginous rod, the stapes. The continued evolution of a particular character past its point of origin is called an *evolutionary transformation series*.

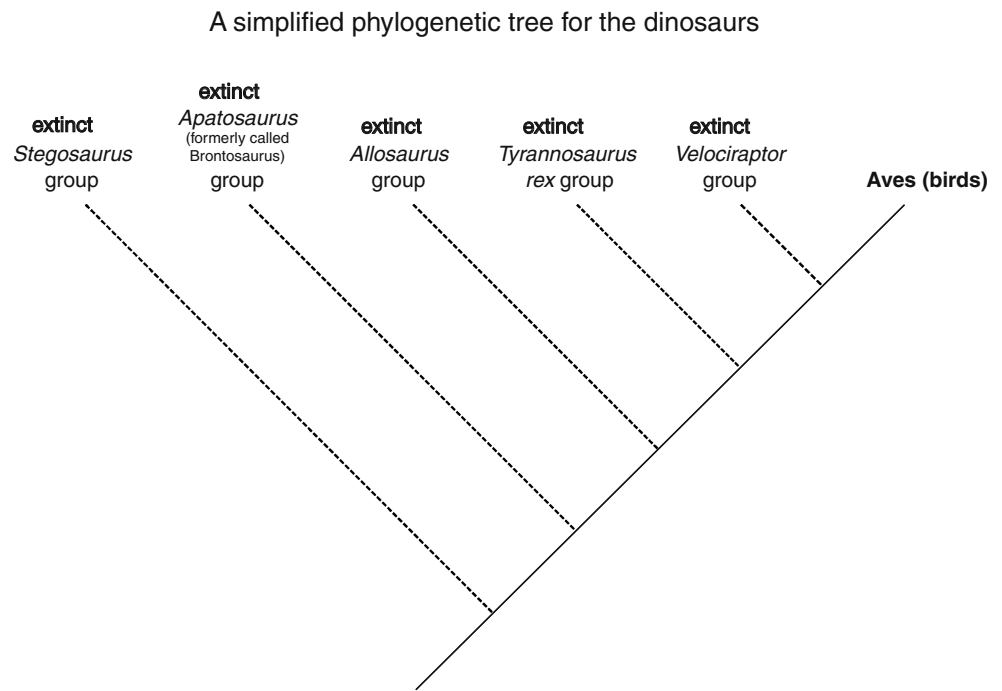
The next thing that students of phylogenetics have to know is how to recognize different kinds of groups of organisms. There are two general types of groups, one “good” and the other “bad”.

Let’s begin with “the good,” a monophyletic group (Fig. 11). The word “monophyletic” is a combination of two Greek words, *monos* (single) and *phyle* (tribe). It was coined by our old friend Ernest Haeckel, who, as you remember, also invented the word phylogeny. A monophy-

letic group includes an ancestor and *all* of its descendants. It is identified by the presence of shared, unique characters (synapomorphies). Each phylogenetic tree contains as many monophyletic groups as there are ancestors. For example, looking at the tree in Fig. 11, we can identify five monophyletic groups, only two of which are shown on Fig. 12 (I’ll leave it up to you to discover the other three).

Now onto “the bad.” The word “paraphyletic” is, once again, a combination of two Geek words, *para* (near) and *phyle* (tribe), so the implication is that the whole tribe is not present (Fig. 13). Paraphyletic groups include an ancestor but *not all* of its descendants. On this hypothetical tree, species C has been eliminated from the group, even though it is a descendant of ancestor 1 just like the rest of the species. Paraphyletic groups are problematic because they mislead us about how characters evolve and how species are related to one another. For example, let’s consider the big tree for the Amniota and highlight the “old” Reptilia, one of the most famous paraphyletic groups (Fig. 14). Even today people still speak about three distinct classes, the reptiles, the birds, and the mammals. When you look at this figure, what is wrong about the class Reptilia, the way it is drawn?

Fig. 7 Actual extinctions. Groups depicted with dotted lines are extinct so all of the genetic, morphological, physiological, ecological, and behavioral traits that are unique to each group have been lost to the biosphere



all of the dinosaur lineages are extinct except for Aves (birds)

Right! In (Fig. 15) Ancestor 2 is the ancestor of all the reptiles but, as highlighted on this figure, the Reptilia does not include all of ancestor 2's descendants; ancestor 4 and the birds have been removed from the group. The only way to make the Reptilia a monophyletic group is to redefine the

term to include crocodiles, turtles, tuataras, lizards, snakes, *and* birds. In the past, birds were not considered to be reptiles because they are warm-blooded (in fact, they were often grouped with mammals because of that trait). But phylogenetic studies have demonstrated that birds are

How to read characters on a phylogenetic tree

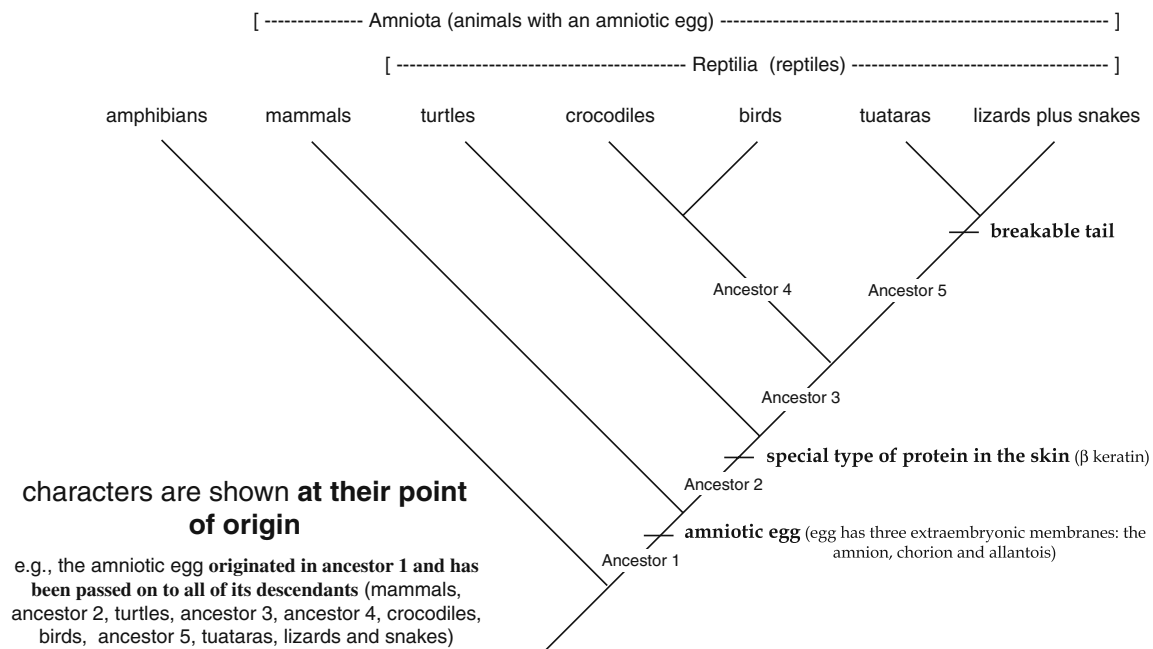


Fig. 8 How to read characters on a phylogenetic tree

Different ways of drawing characters on a phylogenetic tree

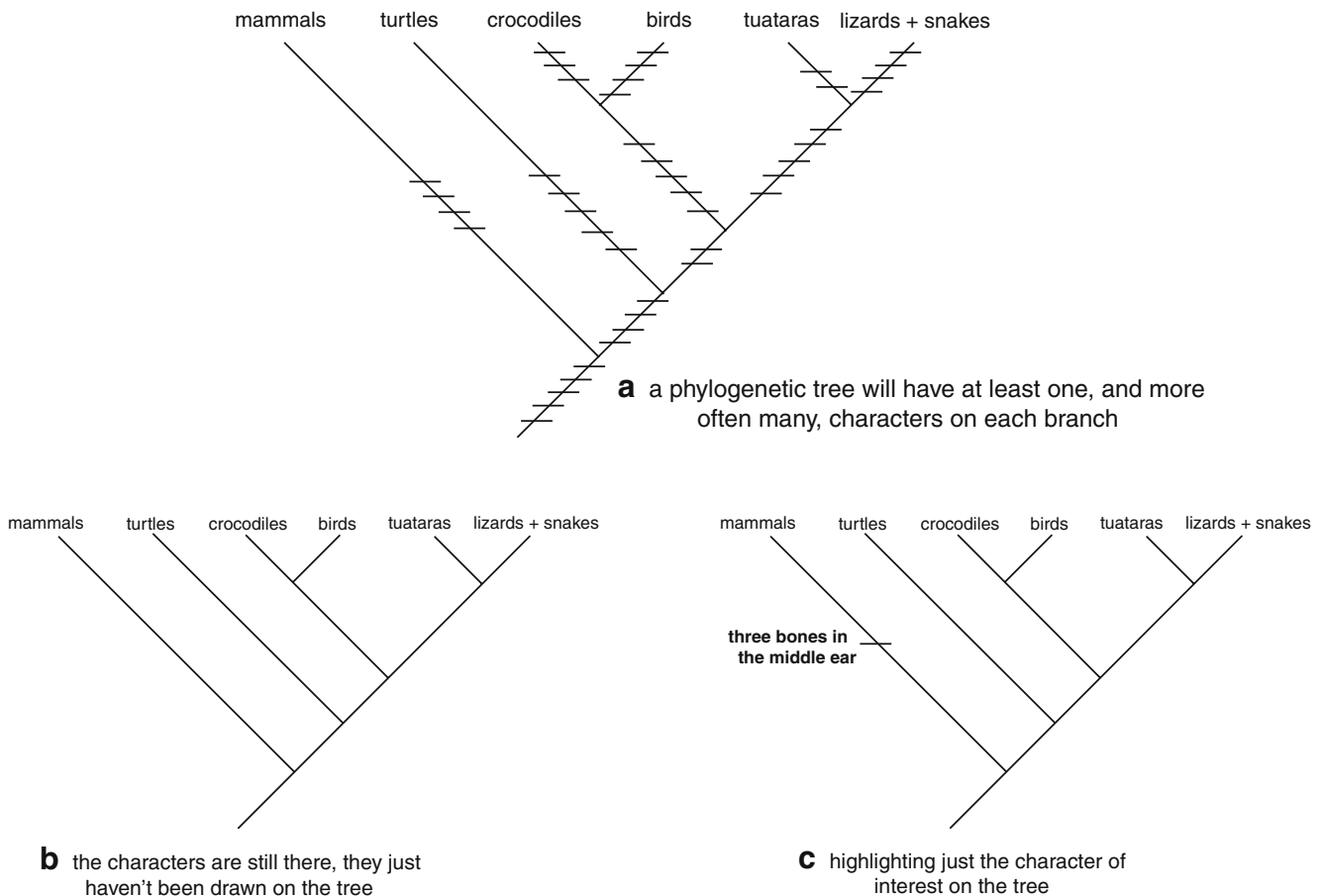


Fig. 9 a–c Representing characters on a phylogenetic tree

indeed reptiles because they share many morphological, behavioral, and molecular characters with other reptilian species in general (synapomorphies originating in ancestor 2; e.g., β keratin), *and* they share many characters with crocodiles in particular (synapomorphies originating in ancestor 4; e.g., holes in the skull just in front of the eyes).

Why is it important to have monophyletic groups? Say you wanted to figure out how red hair appeared in your family. What would be your chances of tracking down your original red-haired ancestor if no records were kept about the union between your great-great-great-great grandfather Sven and his Irish bride Maggie? Missing information creates problems for any research, be it genealogical or evolutionary, and paraphyletic groups are missing information. In evolutionary terms, monophyletic groups are “real” biological units; that is, they are the product of descent with modification (an ancestor and all of its descendants) and as such can be used to study the evolutionary processes that produced them. Paraphyletic groups, on the other hand, are the product of “human

error” arising from incomplete or flawed information (e.g., poor descriptions of characters). Using such groups to study evolutionary processes will direct us along misleading and confusing pathways.

Why do we use phylogenetic trees? There are many ways to answer this question (and many papers/books written about it), but the most general answer is that trees summarize valuable information about the evolution of organisms that allows us to understand them better. For example, here’s the family tree for the Hominoidea, the group that includes us and all of our closest relatives (Fig. 16). When you look at the distribution of characters on this tree you can see that a number of traits we associate only with human beings, such as hunting, infanticide, tool making, self-awareness, and language, originated long before *Homo sapiens*. In other words, human beings are not as unique as you might think. If we want to understand how and why those traits evolved, we must study their expression and function in ourselves and in our relatives. So much information from just one phylogenetic tree!

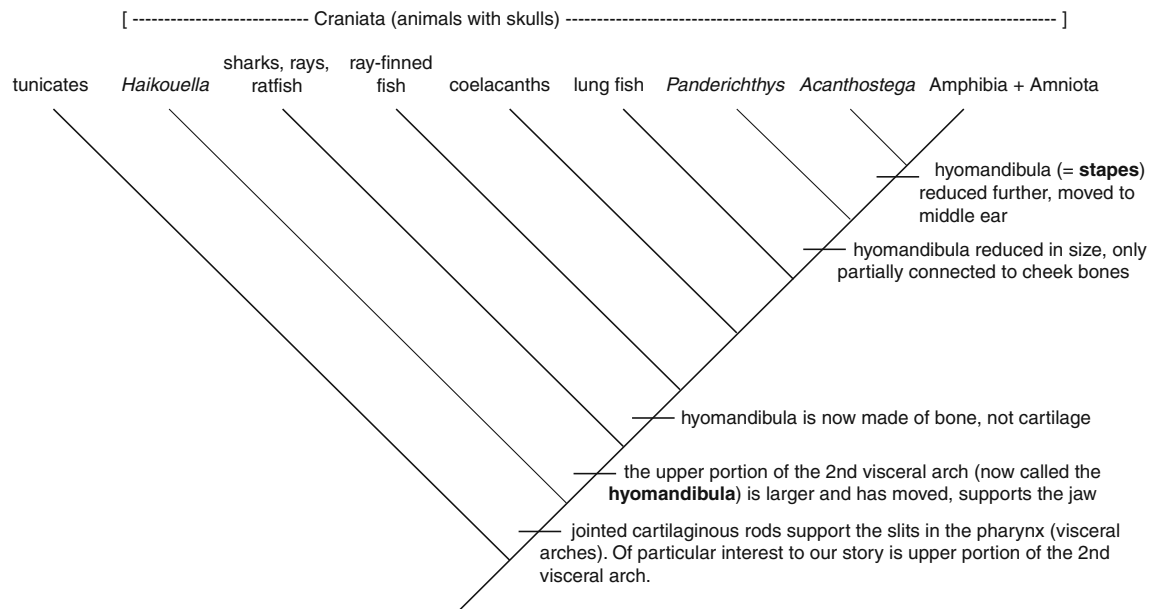
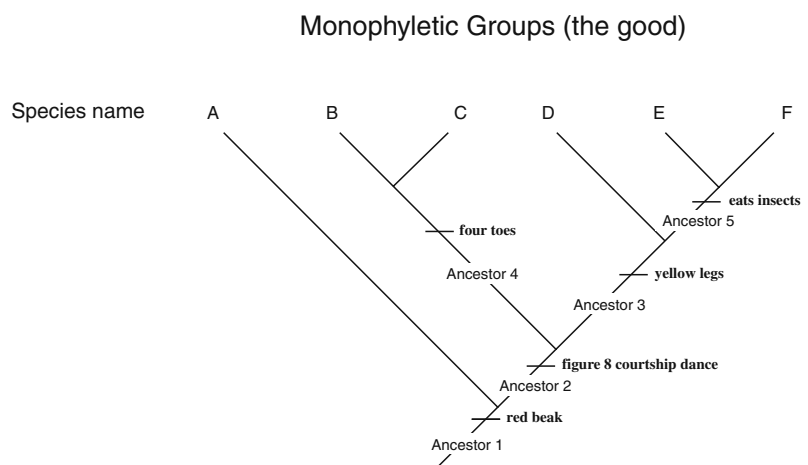


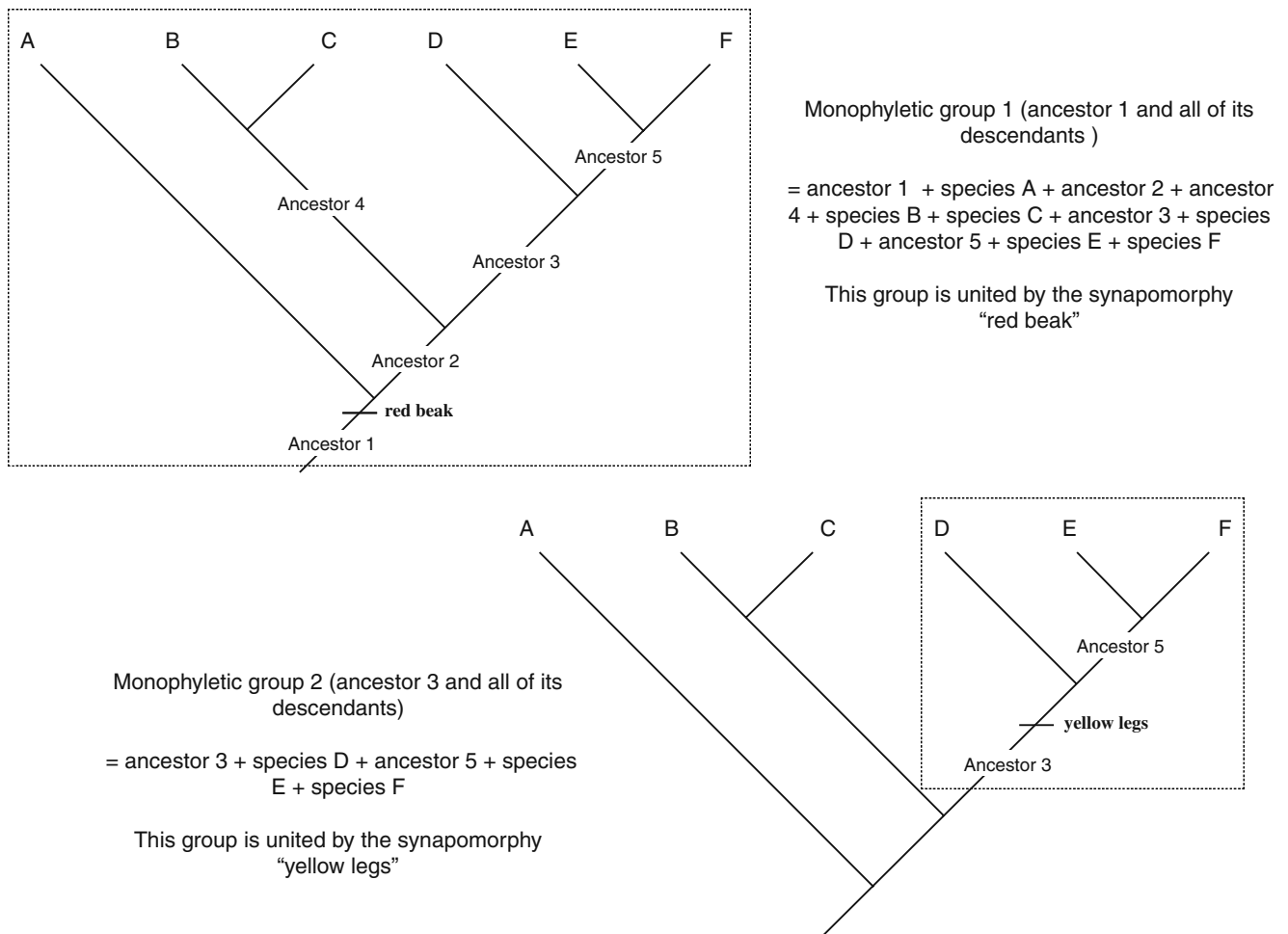
Fig. 10 Synapomorphies are not static; they may continue to evolve. Changes in the character “upper portion of the second visceral arch” [hyomandibula, stapes] are traced on the phylogenetic tree for the Chordata (animals with notochords). Both the story and the phylogenetic tree have been substantially simplified to emphasize

the idea of character origin and modification rather than the finer details of character evolution. *Names in italics* refer to extinct species known from fossils. Line drawings and photographs of various structures and species can be found easily on the web

Fig. 11 Identifying monophyletic groups



There are as many monophyletic groups on this phylogenetic tree as there are ancestors.

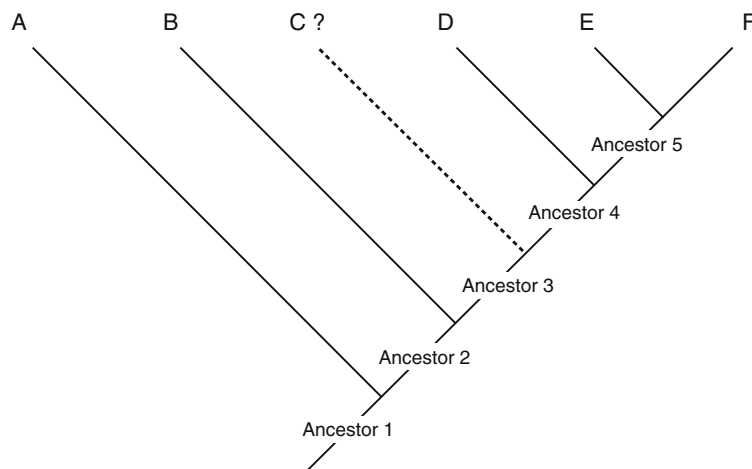


can you identify the other three monophyletic groups on this tree?

Fig. 12 Two of the five monophyletic groups on the hypothetical tree

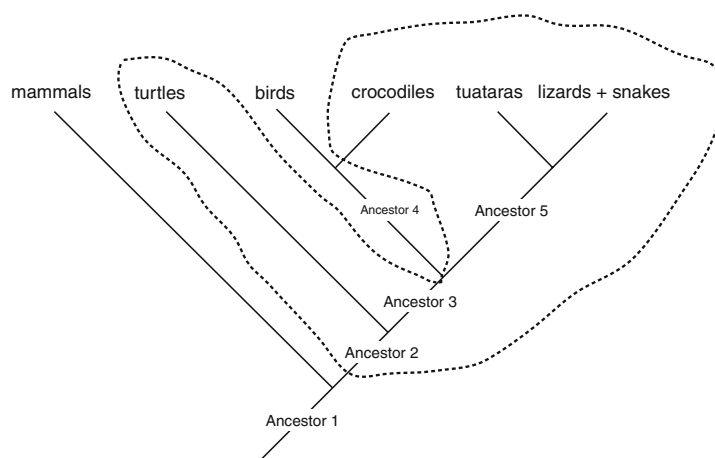
Fig. 13 Identifying paraphyletic groups

Paraphyletic Groups (the bad)



A paraphyletic group includes an ancestor **but not all of its descendants** (e.g., ancestor 1 + species A + ancestor 2 + species B + ancestor 3 + ancestor 4 + species D + ancestor 5 + species E + species F)

What happened to species C?

Fig. 14 The most famous paraphyletic group, the reptilesThe old “class Reptilia”:
the world’s most famous paraphyletic group

The dotted line surrounds taxa that people used to include in the class Reptilia. Even today, many people still think that reptiles are just snakes, lizards, turtles and crocodiles (not many people know about tuataras).

What is wrong with this picture?

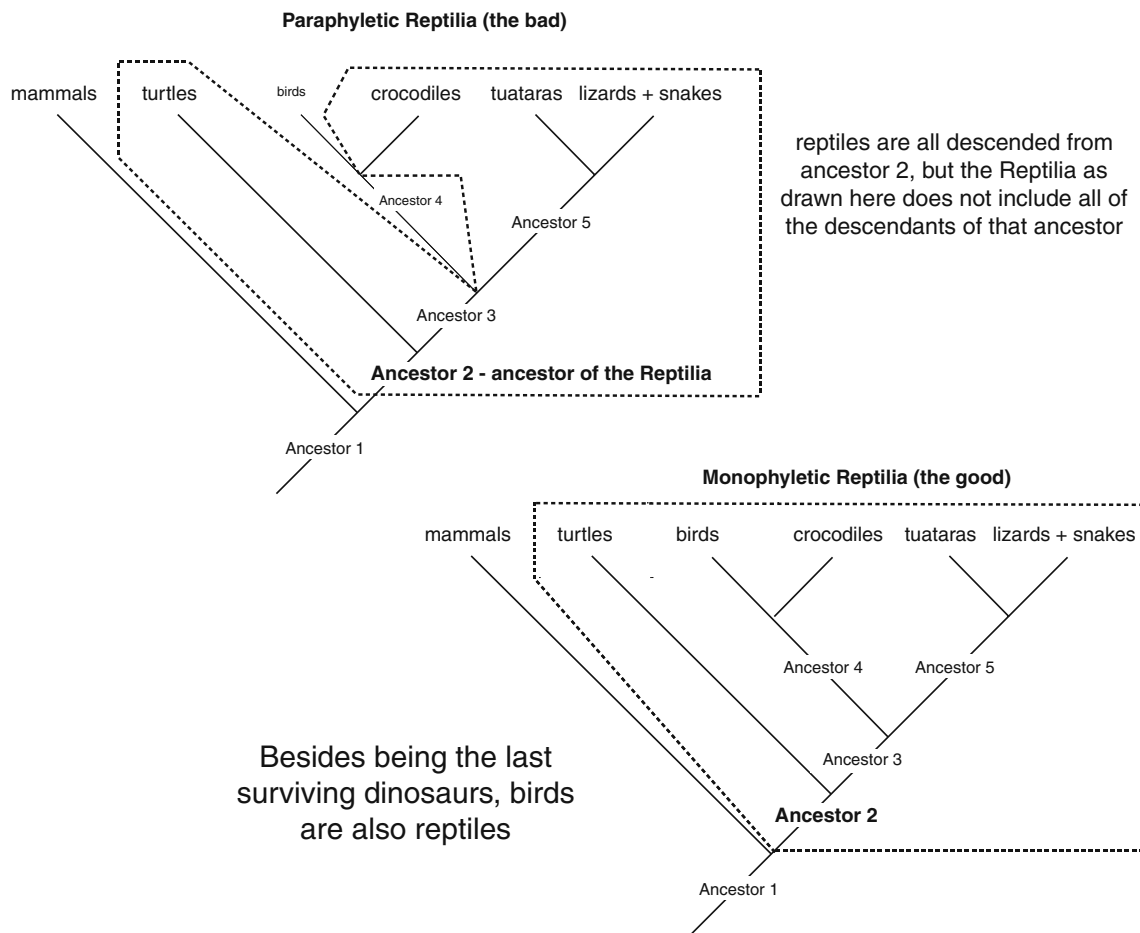
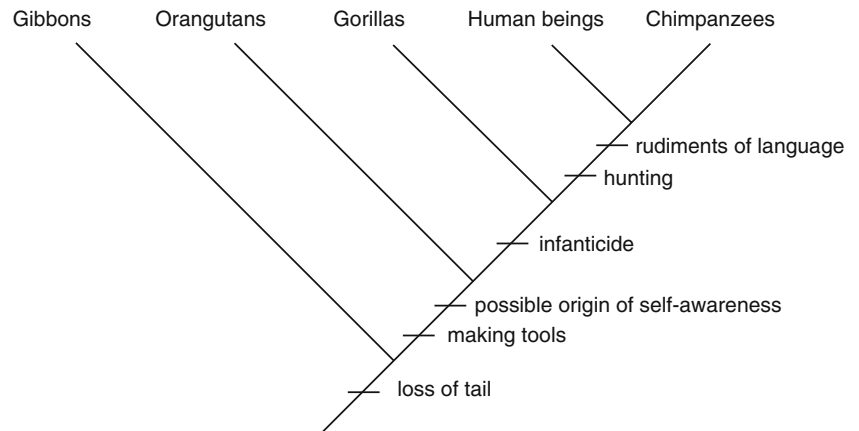


Fig. 15 How to make the Reptilia monophyletic

Fig. 16 Using a phylogeny to study ourselves

Why do we use phylogenetic trees?



The information in a phylogenetic tree can be used to study all aspects of evolution, from character origin and diversification to speciation. That knowledge, in turn, has profound implications for understanding where we came from, what our place is in the biosphere and what we can do to protect that biosphere.