Phylogenetics II: Building trees from organismal data

Introduction to Evolution and Scientific Inquiry
Dr. Stephanie J. Spielman; spielman@rowan.edu

What are all possible trees for three taxa A, B, C?

Two broad approaches to determine the "best" tree

- Either way, we use **homologous characters (traits)**
    - Since early 1990s, almost always **DNA sequences!!!**

- Using **distance** among all sequences (shortest distance = most closely related)

- Using an **optimality criterion**

What are all possible trees for N taxa?

Number of Species	Number of Possible Trees
3	3
4	15
5	105
6	954
7	10,395
8	135,135
9	2,027,025
10	34,459,425
11	654,729,075
12	13,749,310,575 ← **about 14 billion**
13	316,234,143,225 ← **over 300 billion**

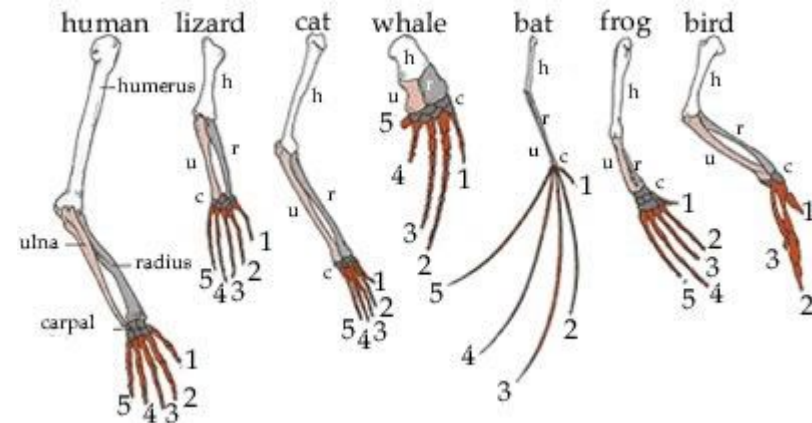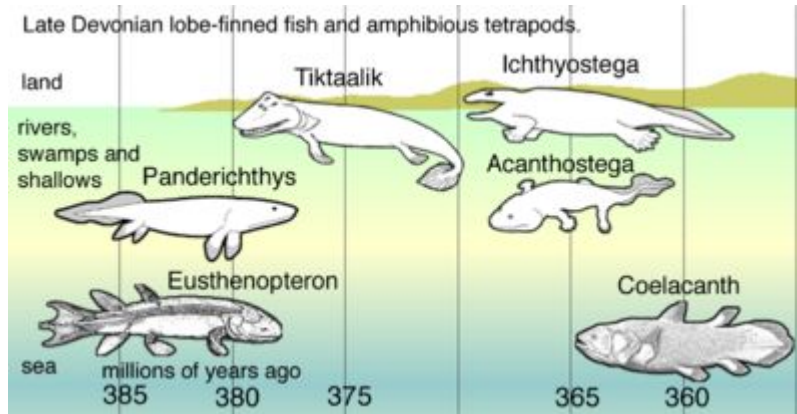# We create trees by tracing change in <u>homologous</u> characters

- **Homology**: traits shared due to **common ancestry**
  - "homologous traits"
  - The trait evolved in the common ancestor, and evolution has "tweaked" the trait as it **diverged** in descendents
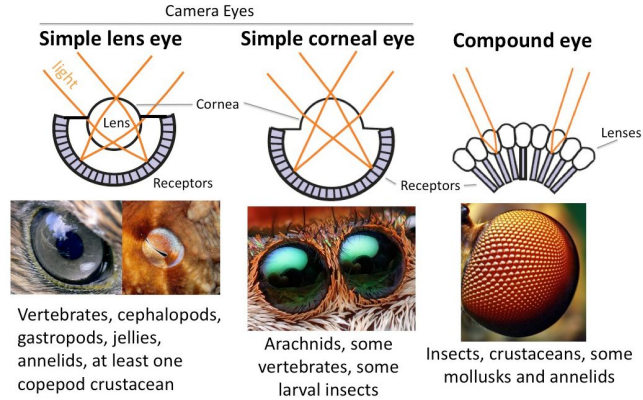

- **Homoplasy**: traits shared due to "acquired" similarity, i.e. **convergent evolution**
  - "analogous traits"
  - The same trait evolves several times independently

Homology of tetrapod limbs

Examples of convergent evolution

Camera Eyes

Simple lens eye

light
Lens
Cornea
Receptors

Vertebrates, cephalopods, gastropods, jellies, annelids, at least one copepod crustacean

Simple corneal eye

Cornea
Receptors

Arachnids, some vertebrates, some larval insects

Compound eye

Lenses

Insects, crustaceans, some mollusks and annelids

Human thumb   Panda thumb

McGraw-Hill Companies, Inc. Permission required for reproduction or display.

Niche	Placental Mammals	Australian Marsupials
Burrower	Mole	Marsupial mole
Anteater	Lesser anteater	Numbat (anteater)
Mouse	Mouse	Marsupial mouse
Climber	Lemur	Spotted cuscus
Glider	Flying squirrel	Flying phalanger
Cat	Ocelot	Tasmanian "tiger cat"
Wolf	Wolf	Tasmanian wolf

shark   ichthyosaur   dolphin

Convergent evolution, phylogenetically
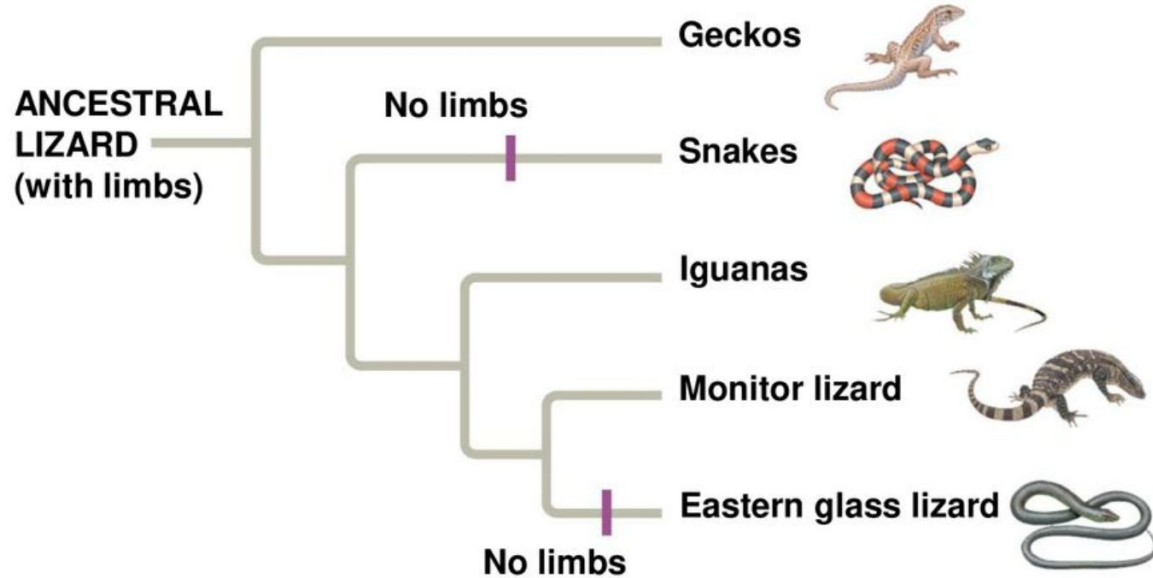
Why homologous characters are key
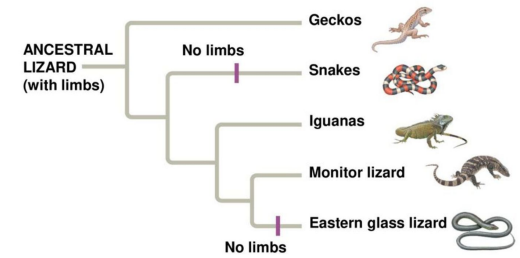
Building trees with homologous characters: **organizing data into a character matrix**

Rows are species/groups of organisms

Columns are trait values for HOMOLOGOUS characters

*amniotic egg:
an egg in which the
embryo is surrounded
by the moisture-retaining
amnion membrane

**post-orbital fenestrae:
holes in the skull
behind the eye

|	Vertebrae?	Bony skeleton?	Four limbs?	Amniotic egg?*	Hair?	Two post-orbital fenestrae?**
Sharks and relatives	YES	no	no	no	no	no
Ray-finned fishes	YES	YES	no	no	no	no
Amphibians	YES	YES	YES	no	no	no
Primates	YES	YES	YES	YES	YES	no
Rodents and rabbits	YES	YES	YES	YES	YES	no
Crocodiles and relatives	YES	YES	YES	YES	no	YES
Dinosaurs and birds	YES	YES	YES	YES	no	YES

Making trees from DNA sequences

If the character matrix is DNA sequences, it is called a **sequence alignment**
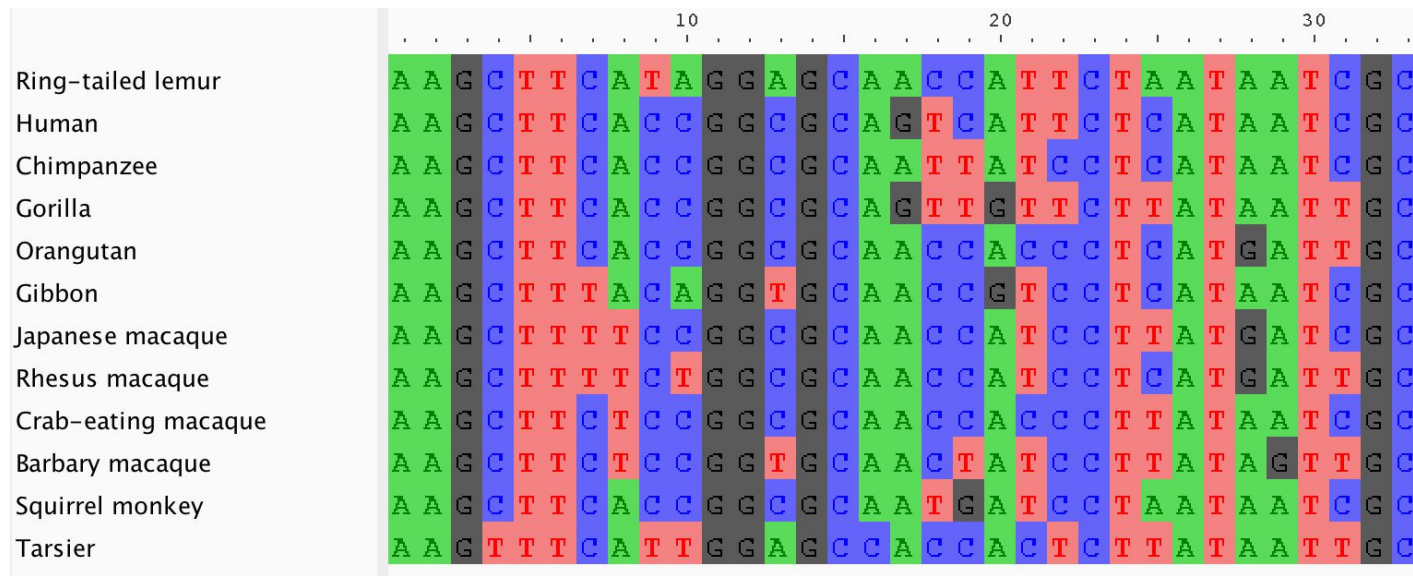
From DNA sequence to phylogeny
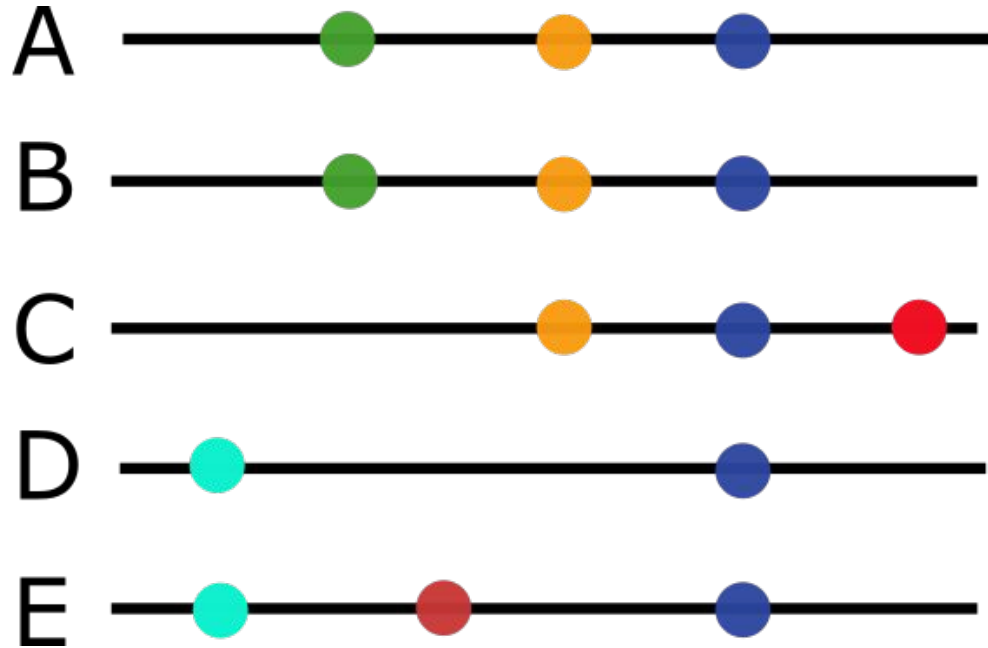
From DNA sequence to phylogeny
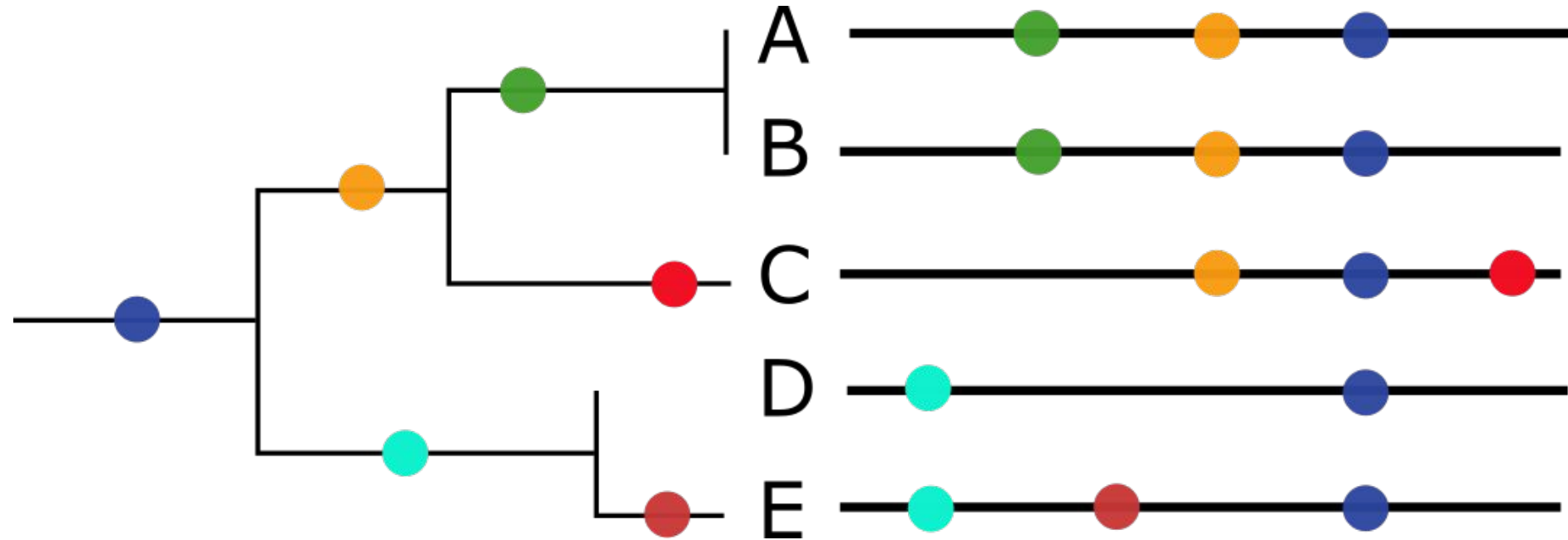
Optimality criterion = a measurement of being optimal (the best)

- Find the tree with the **best value** (optimality) of some measurement (criterion) that tells us if the tree is a **good fit to the data**

- Good fit to the data = the tree and data match really well

- (Note, there are other ways also, but this is the modern-day standard)

Types of optimality criterion

- Parsimony
    - The tree with the fewest steps/evolutionary changes is the best tree
    - **We will learn this one**
    - For all possible trees, the one with the fewest number of changes is the "best"

- Some kind of complicated statistic
    - "Maximum Likelihood"
    - "Bayesian Posterior Probability"
    - **By FAR the most commonly-used approaches in modern-day phylogenetic research**
    - For all possible trees, the one with the highest PROBABILITY is the "best"

Major question: Is evolution parsimonious??

PROBABLY NOT.

https://twitter.com/RebeccaRHelm/status/1245810190601072653?s=20

Let's find the best tree under <u>parsimony</u>

```
Fish:     GCGT
Bear:     CCTG
Lizard:   CCAG
ancestor: GCGT
```

Position 2 is <u>constant</u>,
1, 3, 4 are <u>variable</u>

The tree with the **lowest tree length** is the best tree. Tree Length = total number of changes along the tree, <u>summed across characters (columns in alignment)</u>. **TREE LENGTH IS OUR OPTIMALITY CRITERION.**

# Once we have a tree, we can study evolution of traits

Bear   **Lizard**   **Fish**
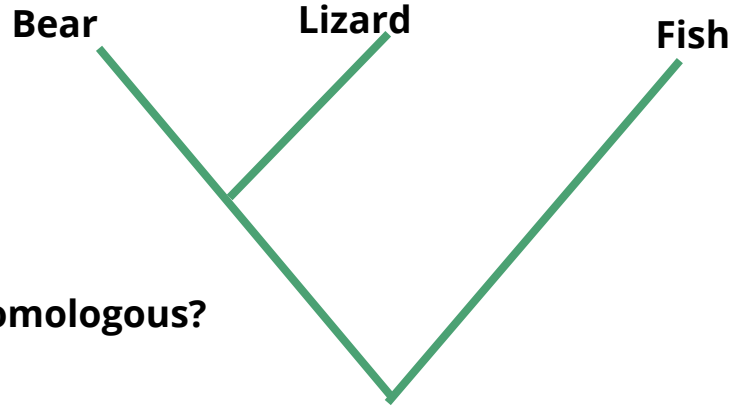
Which traits are likely homologous? Homoplasious?

	Four limbs	Lives on land	Eats insects
Bear	Yes	Yes	No
Lizard	Yes	Yes	Yes
Fish	No	No	Yes

Tree-thinking about traits/groups of species

Which is a true evolutionary group?

- **Animals with a four-chambered heart**
  - **Birds and mammals**
- **Birds ("Aves")**
- **Reptiles**
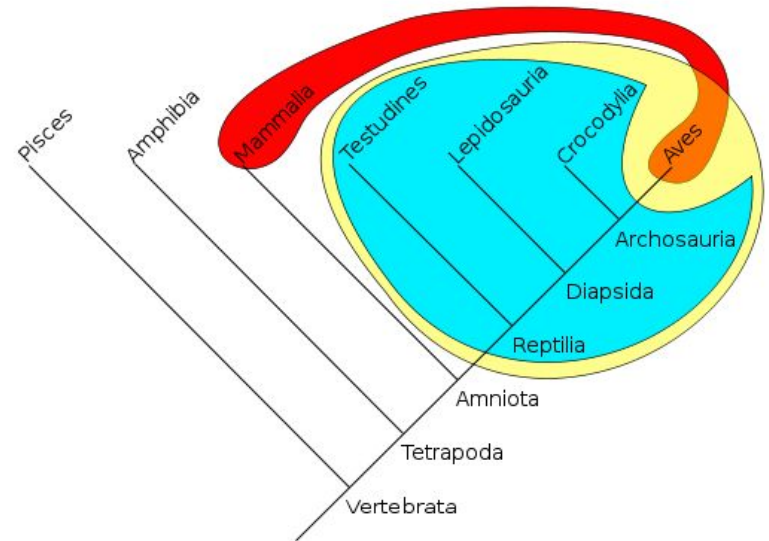 - **lizards, turtles, snakes, crocodiles**
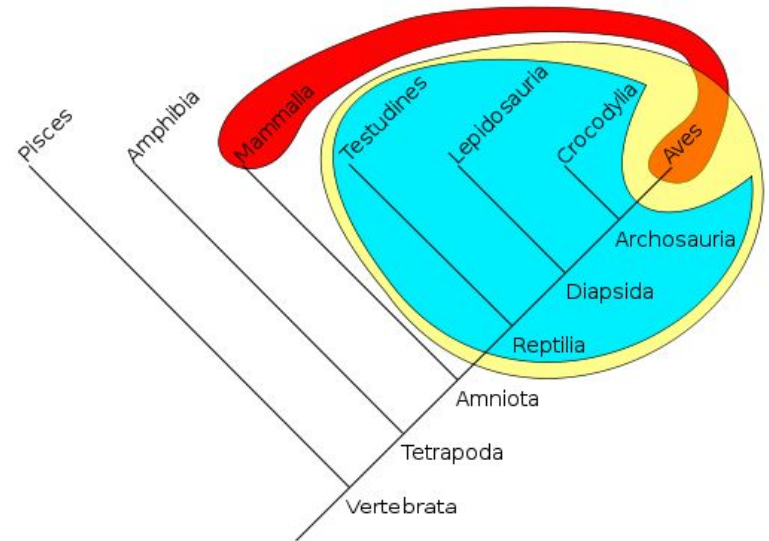
Tree-thinking about traits/groups of species
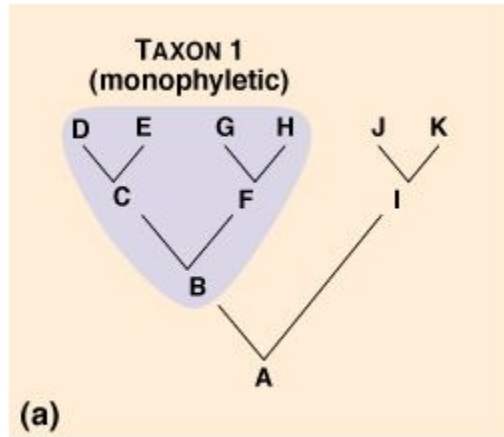
Which is a true evolutionary group?

- **Animals with a four-chambered heart**
  - **Birds and mammals**
- **Birds ("Aves")**
- **Reptiles**
  - **lizards, turtles, snakes, crocodiles**

Describing groups in a phylogenetic context
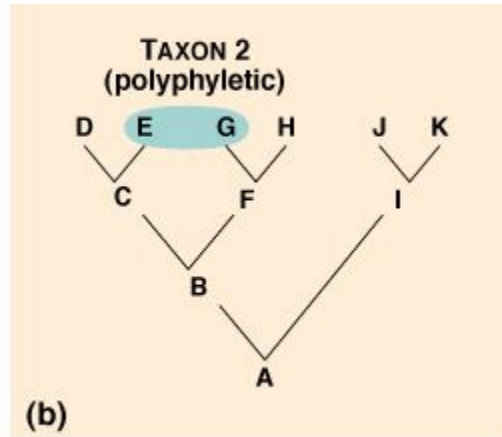
Letters at nodes represent *labeled ancestors*

TAXON 1 (monophyletic)

TAXON 2 (polyphyletic)

TAXON 3 (paraphyletic)

©1999 Addison Wesley Longman, Inc.

CONVERGENT trait
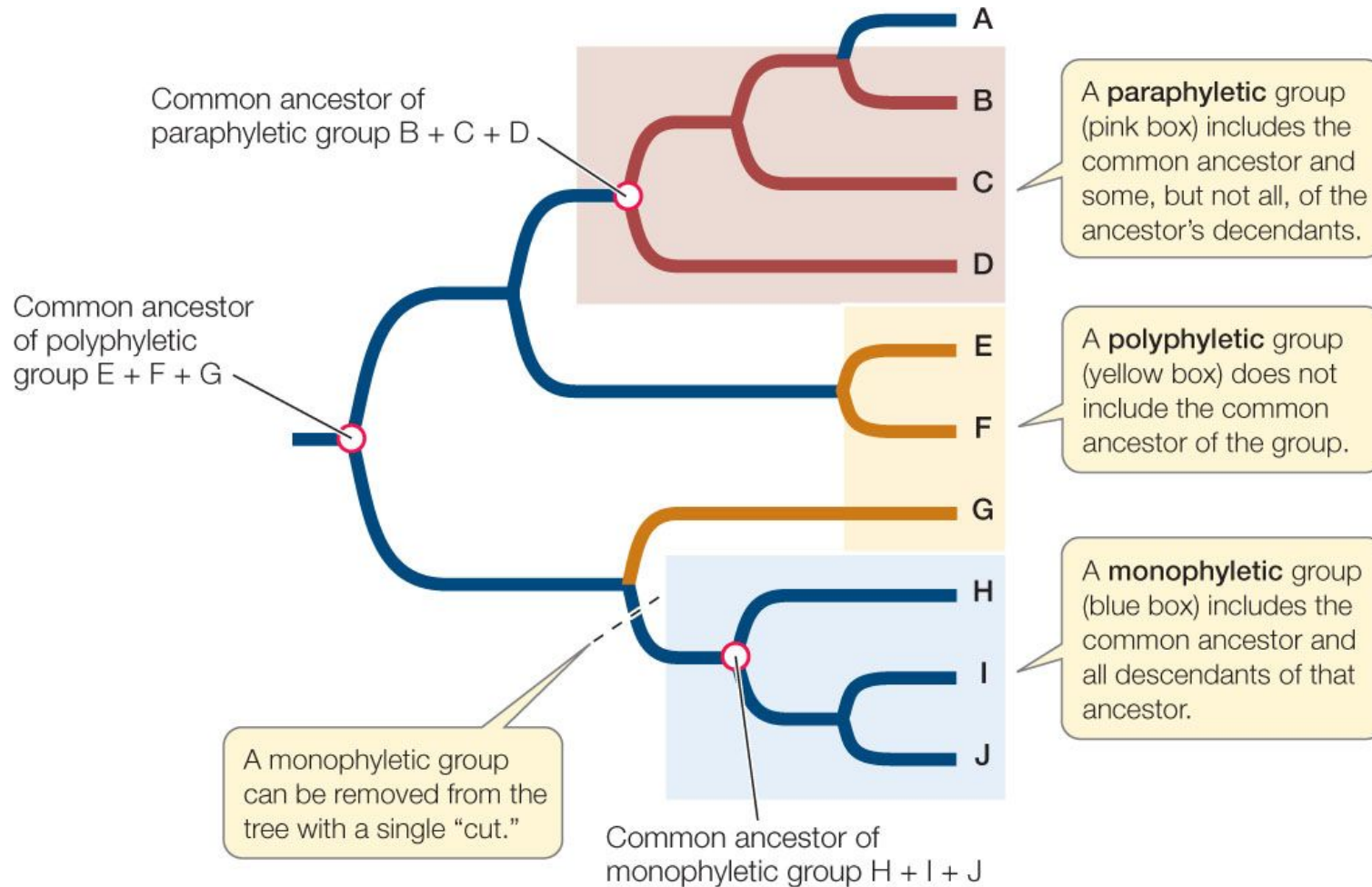
Homologous trait with all descendants inheriting same trait
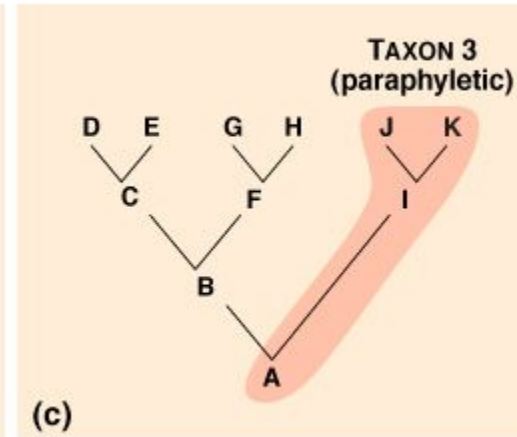
Homologous trait with some descendants inheriting trait, rest inheriting *modification*

Common ancestor of paraphyletic group B + C + D

Common ancestor of polyphyletic group E + F + G

A **paraphyletic** group (pink box) includes the common ancestor and some, but not all, of the ancestor's decendants.

A **polyphyletic** group (yellow box) does not include the common ancestor of the group.

A **monophyletic** group (blue box) includes the common ancestor and all descendants of that ancestor.

A monophyletic group can be removed from the tree with a single "cut."

Common ancestor of monophyletic group H + I + J

"No limbs" is **polyphyletic** (ancestor NOT in group - it has limbs!)
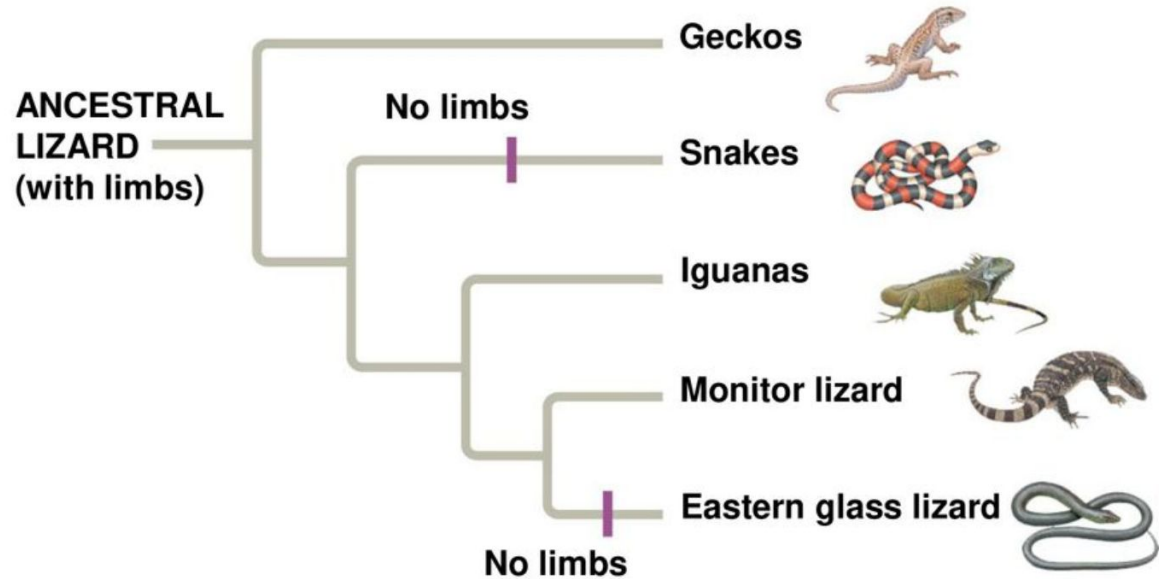
Marine mammals show convergence. They are polyphyletic

Cladogram for Marine mammals (simplified)

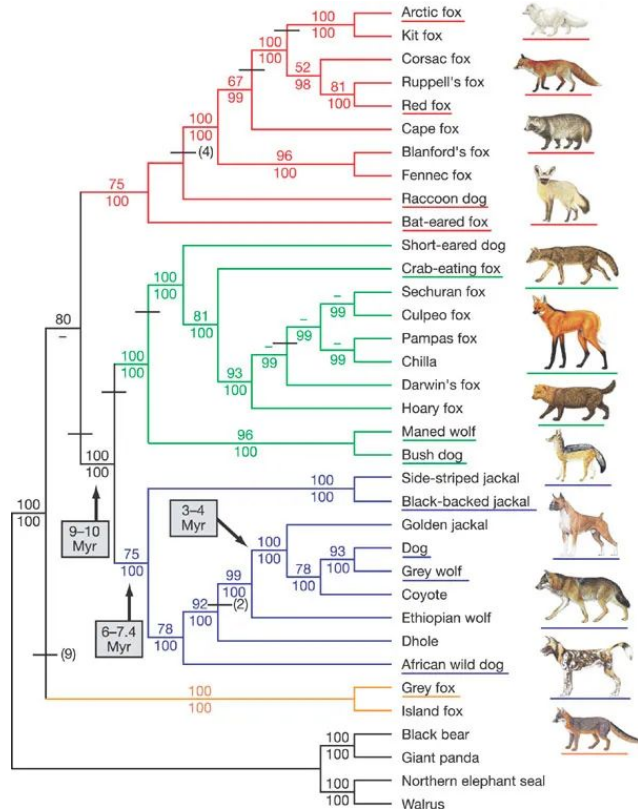Marine mammals is a polyphyletic group
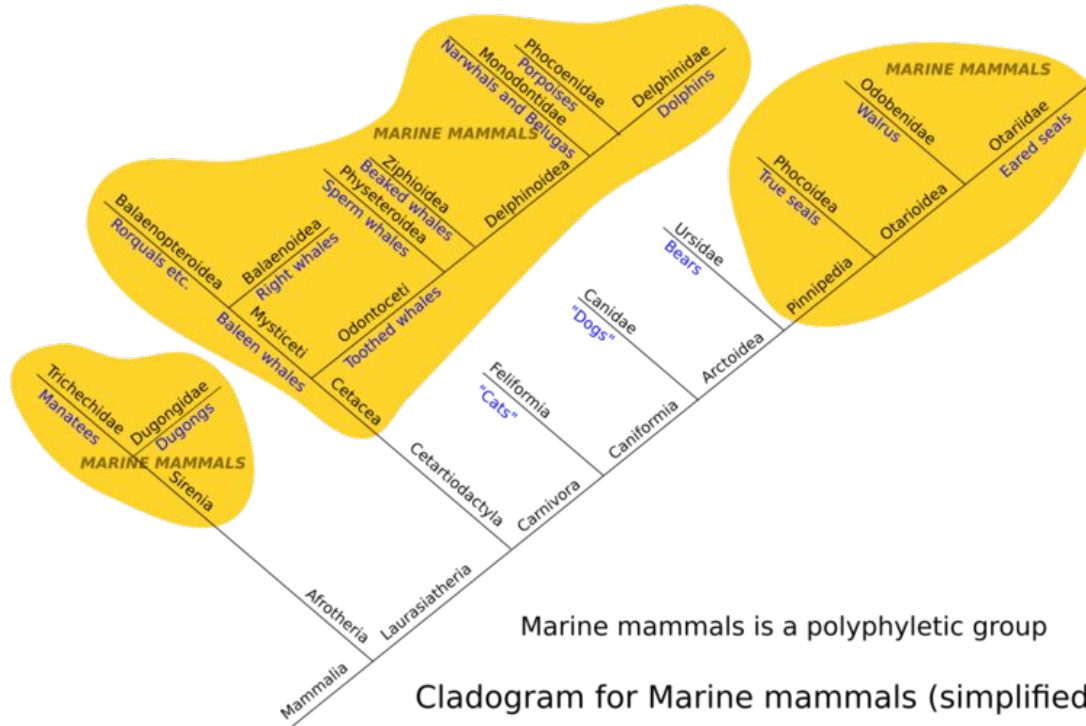
Birds are MONOphyletic. So are crocodilians. Dinosaurs are PARAphyletic

Lizards and snakes

Crocodilians

Common ancestor of crocodilians, dinosaurs, and birds

Ornithischian dinosaurs

Saurischian dinosaurs

Birds
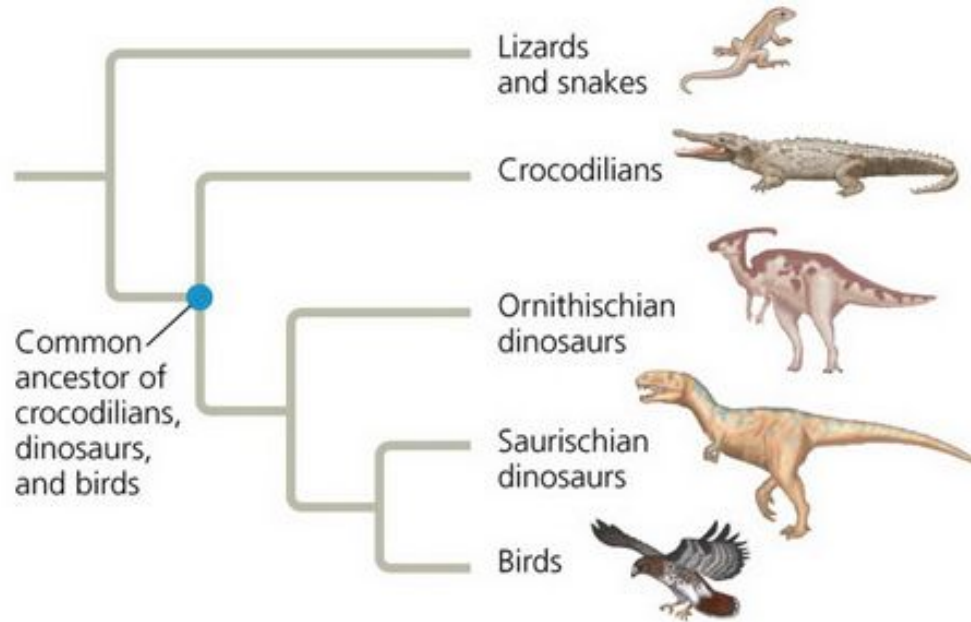
Dogs and their close relatives

"The phylogenetic tree is based on ~15 kb of exon and intron sequence (see text). Branch colours identify the red-fox-like clade (red), the South American clade (green), the wolf-like clade (blue) and the grey and island fox clade (orange). The tree shown was constructed using maximum parsimony as the optimality criterion and is the single most parsimonious tree.

"Bootstrap values and bayesian posterior probability values are listed above and below the internodes, respectively; dashes indicate bootstrap values below 50% or bayesian posterior probability values below 95%.
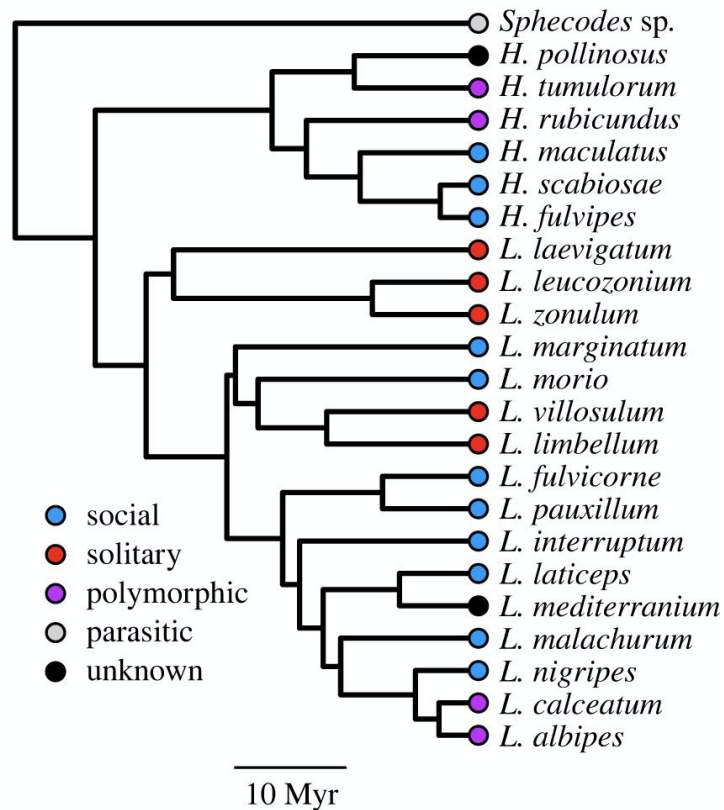
"Underlined species names are represented with corresponding illustrations."

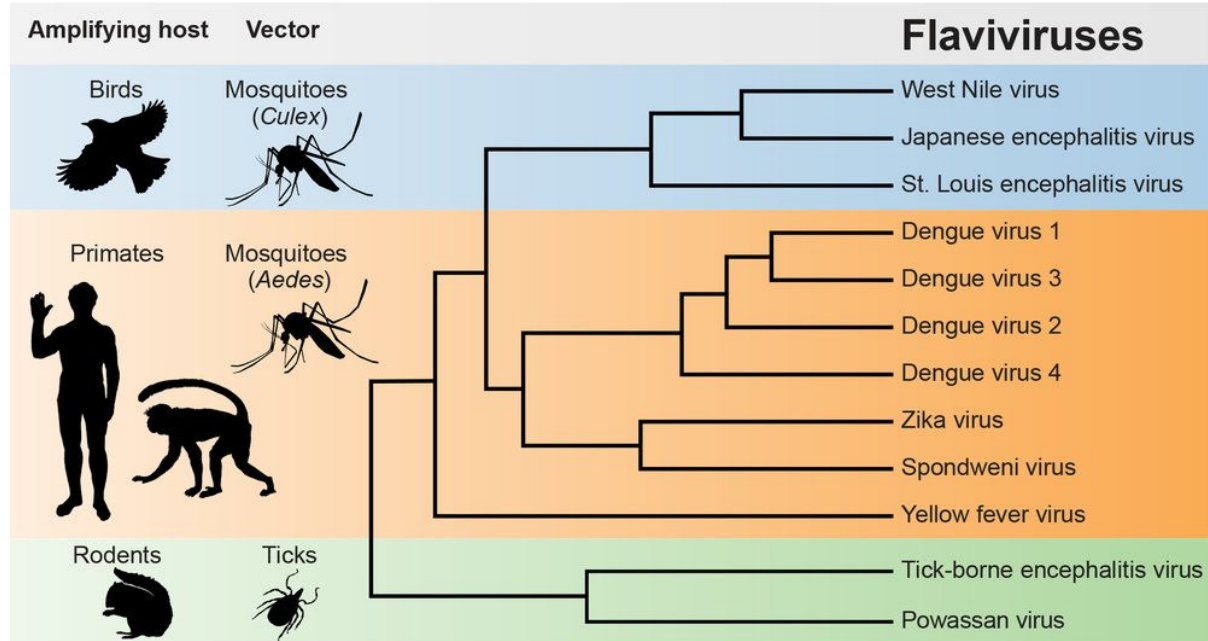https://www.nature.com/articles/nature04338

Evolution of bee behavior

Are of of these groups **monophyletic?**

- Social?
- Solitary?
- Polymorphic?
- Parasitic?

How many **evolutionary changes** have occurred for this trait?

Flavivirus hosts?

(This is a **cladogram**: branch lengths are meaningless.)