

# Bio 1M: Phylogeny and the history of life (complete)

## 1 Phylogeny

- *Bioskills are in the back of your book*
- Phylogeny is the evolutionary history of a group of organisms
- Based on the idea that organisms are related by evolution
- Understanding these relationships is critical to our understanding of both evolution, and how biological processes work

### Phylogenetic trees

- A **phylogenetic tree** is a model of how a group of organisms descended from a common ancestor
- The model consists of **nodes**, where groups split, **branches** where evolution occurs, and **tips** representing observed **taxa** which are the endpoints of the process we are trying to model.
- A point where more than two branches diverge is called a *polytomy*.
  - Polytomies mean we don't know which event happened first.

### Monophyletic group

- A **monophyletic group** is a group *defined by* a single common ancestor
  - All descendants of the ancestor must be in the group
- Monophyletic groups can also be called **clades** or **taxa**.
- As biologists, we should try to think in terms of clades
  - Are flying vertebrates a clade?
  - What are some prominent groups that are not clades?
    - \* **Answer:** apes, reptiles, trees

### Sister taxa

- Sister taxa can be a useful way of thinking about trees
  - two taxa that share a common node
  - You need to take the whole taxon
- E.g., sisters of: *Homo sapiens*; *Homo erectus*; humans?
  - **Answer:** *Homo neanderthalensis*; *Hn* and *Hs*; *Paranthropus*

## Reading phylogenetic trees

- The tree indicates the pattern of branching of **lineages** (evolving lines)
- Tips are *assumed* by the model to be monophyletic
- A tree is a model of how evolution occurred
  - Trees that correspond to the same model are considered equivalent

## Order of species

- Except for the branching pattern, we don't interpret anything about the order of a tree
  - No species are higher or lower than others, according to the tree
  - We've all been evolving for the same amount of time
- How do we judge how closely related two organisms are, according to a tree?
  - **Answer:** By looking for their common ancestor

### 1.1 Constructing phylogenetic trees

- First, measure **characteristics** (or **characters**) of the taxa of interest – i.e., anything that seems useful to measure
  - Can be **morphological** (i.e., physical) or genetic
- Then infer (make an educated guess about) the phylogenetic relationships.
  - **Phenetic** approaches use measures of distance between organisms
  - **Cladistic** approaches are based on modeling how evolution occurs on the tree

## Morphological vs. genetic characteristics

- Should we use morphological or genetic characteristics to infer phylogenies?
  - We usually have more information from genetic characteristics, and this information is easier to measure precisely
- When should we use morphological characteristics?
  - **Answer:** When genetic information is not available
    - \* **Answer:** Most fossil organisms
  - **Comment:** Some viruses evolve so fast that morphological characteristics can be more stable than genetic characteristics

## Phenetic vs. cladistic analysis

- Cladistic analysis makes use of the phylogenetic model of organisms evolving from each other to infer phylogenies
- Phenetic analysis ignores the phylogenetic model of organisms evolving from each other while inferring phylogenies
- Which is better?
  - **Answer:** Cladistic analysis
- When should we use the other one?
  - **Answer:** When our data are only distances, with no characters available – e.g., certain molecular techniques
  - **Answer:** When we don't have enough baseline information

## Synapomorphies

- Classical cladistic analysis is based on **synapomorphies** – shared, **derived** characters
  - as evidence that two taxa are related
- Why do we focus on derived characters?
  - **Answer:** These are things that evolved in the relevant context
  - **Answer:** So they are the things our model (the phylogenies) must explain

## Flight as evidence

- Oaks and fish don't fly, but birds do
  - **Answer:** We're pretty sure that the common ancestor didn't fly
  - **Answer:** No evidence (any model will have flight evolve once)
  - **Answer:** Flight evolved once in either case
- Mosquitoes and birds fly, but fish don't
  - **Answer:** Evidence (flight might have evolved once or twice on this tree)
  - **Answer:** But this evidence points in the wrong direction
    - \* **Answer:** This is why we combine evidence from different sources

## What characters are derived?

- Derived compared to what?
  - The common ancestor (characteristics of the common ancestor are called **basal** or **ancestral** characters).
    - \* What if our flight example was ostriches, hawks, and sparrows?
    - \* **Answer:** No evidence for similarity between hawks and sparrows
    - \* **Answer:** We believe the common ancestor could fly, so there is no shared, derived characteristic
  - A **derived** character is a character not shared by the common ancestor of the group that we are currently thinking about
  - One way to think about cladistic vs. phenetic analysis is that phenetic analysis treats derived and basal characters equally

## Inferring the common ancestor

- We want to know what the common ancestor was like, so we can tell which characters are derived (as opposed to basal)
- This can be done sometimes by common sense
- Inferring the common ancestor statistically is difficult, for technical reasons
  - We can make use of an **outgroup** to study a group. An outgroup is an organism closely related to, but outside, the group being studied.
  - We assume that the the **root**, or beginning, of the tree is where the outgroup branches from the group

## Confusing the phylogeny

- Two species may have the same trait because the trait evolved twice independently – **convergent evolution**
  - **Answer:** Flight, trees, dolphins and ichthyosaurs
- An organism may lack a character that its ancestors had – **secondary loss**
  - This can be even more confusing
- Similarities that are not homologies (ie., not due to common ancestry) are called **homoplasies**

## Parsimony

- Classical cladistic analysis is based on searching for the tree that can explain the observed data most **parsimoniously** – with the fewest number of changes necessary.
- How do we address the problem of convergent evolution and homoplasy?
  - **Answer:** Make use of many different characteristics, when possible
  - **Answer:** It may also help to use many different taxa
- *Comment:* Modern approaches that use genetic data may use more sophisticated approaches, rather than simple parsimony

## 1.2 Example: the evolution of whales

- Whales seem, on the surface, to be pretty different from other mammals, including the closest group, the artiodactyls
- Early phenetic analyses showed whales in their own group (sisters to artiodactyls)
  - Whales lack the distinctive artiodactyl ankle structure (surprise!)
- What about cladistic analysis?

### Activity: Tree-construction analysis

- Hippos have derived characters A, B, C, D, Z
- Cows have derived character A, Z
- Whales have derived characters B, C, D, E, F, G, H, I, Z
- Calculate distances, and make a phenetic tree
- List synapomorphies, and make a cladistic tree

### Tree discussion

- What is the effect of universal characteristics (like Z, above)?
  - **Answer:** None
- What is the effect of unique characteristics (like E, F, G, H, I)?
  - **Answer:** They affect phenetic trees, but not cladistic trees
  - **Answer:** As long as we are sure that they are derived!
- Why might whales have more derived characters than the other species?
  - **Answer:** Because they have had to adapt more since moving to the water

## Confirmation

- Intermediate forms between hippo-like animals and whales
- Discovered *after* they were predicted from DNA evidence!

## How to make a phenetic tree

- 1. Join the two open nodes that are closest to each other
  - Shortest distance
- 2. They are no longer active. Make a new node halfway between them
  - Average the traits
  - This step is conceptually simple and practically tricky
  - Comment: This is why we have computers
- 3. Go back to step 1

## How to make a phenetic tree in this course

- 1. Join the two open nodes that are closest to each other
  - Shortest distance
- Answer: 2. *Stop!* You're done.
  - Answer: I will only give you three species
  - Answer: Seriously.
- Comment: I'm here to teach you concepts, not to teach you how to be a really bad computer.

## Lessons

- Phenetic analysis uses less information
- To do a cladistic analysis, inferences about the common ancestor can be crucial
  - What if we thought the artiodactyl ancestor was aquatic?
    - \* E, F, G, H and I could all be basal characters
    - \* Answer: We would conclude that cows are closely related to hippos

## Genetic vs. morphological

- Why was the genetic analysis more effective than the morphological?
  - Answer: It can be hard to tell which traits are derived
    - \* Answer: This can also be a problem with the genetic analysis
  - Answer: Genetic analyses typically allow us to analyze more traits

## Phylogenetic trees are approximations!

- Phylogenetic trees are tremendously useful and powerful tools for organizing, understanding and analyzing biological data
- But they have limitations:
  - The true history of life cannot really be summarized by a tree
    - \* Sex, other forms of combination or gene transfer
  - Trees constructed by humans are not necessarily even the best approximations to the true history of life
  - Our guesses often change over time

## 2 The history of life

### 2.1 The shape of the tree

- Big recent innovations have told us a lot about the history of life
  - Detailed genetic information
  - Sophisticated analysis techniques
  - Electronic computers

### Three domains

- Bacteria
  - no nuclei; mostly small
  - most of the micro-organisms you see
- Archaea
  - no nuclei; mostly small
  - rarer, or live in more extreme environments
- Eukarya
  - large, nucleated cells

### Eukaryotes

- Seem to be sisters with archaea, according to most of our key genes
- Characterized by **nuclei** and **mitochondria**
- Mitochondria came later, and are captured bacteria!
- Where did the nucleus come from?
  - Nobody knows

## The web of life

- The tree is only part of the story!
  - If genes (or even whole bacteria) can be transferred, life is not really a tree
- Reuniting can create new species
  - Hybridization
  - Allopolyploidy
    - \* **Answer:** Polyploids arising from different (allo-) species
- Wolf-coyote example

## Five kingdoms

- Not really a good way to describe the evolution of life
- So why has this idea persisted for so long?
  - **Answer:** It matches the way the world *looks* – animals, plants and fungi dominate *visible* life
    - \* **Answer:** Some branches have organisms that are much smaller, or just have much fewer organisms
  - **Answer:** Archaea and bacteria are difficult to tell apart even with a microscope

## 2.2 The fossil record

- A **fossil** is a physical trace of an organism from the past
  - Intact fossils retain their form and substance
  - Compression fossils are squashed into a thin film
  - Cast fossils occur when the decomposing piece is replaced by minerals different from the surrounding ones
  - Permineralized fossils occur if minerals infiltrate cells as they are decomposing
- *Comment:* Does any of this seem likely?

## Biases in the fossil record

- Scientists have learned a lot from studying fossils, but care is needed. Very few things fossilize, and some things are much more likely to fossilize than others, for example:
  - Things that live in swampy areas, or underground (**Habitat bias**)
  - Hard things, or hard parts of things (**Taxonomic bias**)
  - Things that lived more recently have had less time to be destroyed, or to be buried too deep for recovery (**Temporal bias**)
  - Things that are more abundant have more chances to be preserved (**Abundance bias**)



## Accounting for biases

- Because the fossil record is biased, scientists must be very careful about making inferences from what they see
  - Just because you don't see it, doesn't mean that it wasn't there
  - Just because you see a lot, doesn't mean that there were a lot (relatively speaking)
- Trying to figure out what happened based on what kind of fossils we happen to find is tricky
- Orchid phylogeny

## 2.3 Putting the timeline together

- Dates of specimens can be inferred using radioactive isotopes (harder than it sounds)
- Geologic inferences can be made about the relative age of different things (for example, if one was deposited on top of another).
- Molecular clocks are based on inferences about how fast things are evolving
- These techniques are complicated, and making the timeline is a difficult and exciting puzzle

## 3 Processes of diversification

- Diversity sometimes arises gradually, and sometimes dramatically (**radiation** events)
- Species (and families, and bigger groups) sometimes disappear gradually, and sometimes dramatically (**mass extinction** events)

### 3.1 Adaptive radiations

- An **adaptive radiation** occurs when a single lineage produces many descendant species, in a short period of time, that make their living in a variety of different ways
- Triggered by opportunity, either in the environment, or because of the evolution of the organisms themselves

## Triggers for adaptive radiation

- Ecological opportunity
  - An organism arrives in an area with no similar organisms
  - A group of competing species is driven extinct (or nearly extinct) by some other cause
- Morphological innovation: an organism comes up with a good, new idea
  - Legs in tetrapods
  - New body plans in the Cambrian explosion
  - Multi-cellularity
- Co-evolution: the evolution of one group creates new niches for another group, and vice versa
  - The insects and the flowering plants

## Colonization

- In many different cases, a new kind of organism has colonized an area, and diversified rapidly
- This is an opportunity for studying how evolution can go:
  - What is the same, and what is different, between similar cases?
  - Finches
  - Silverswords in Hawaii
  - Anolis lizards

## Competition

- Mammals did very well after the dinosaurs went extinct
- We have seen major tree extinctions due to introduced diseases in the last century
  - **Answer:** This opens up ecological niches
  - **Answer:** Species may diverge to fill these niches
  - **Answer:** Or spread around the world to fill them

## Morphological innovation

- A new adaptive mutation can open up further possibilities for adaptation
  - **Hox gene** mutations allowed early animals to develop complex body plans
  - The arthropod body plan
    - \* insects, arachnids, crustaceans ...
  - The tetrapod body plan
    - \* reptiles, mammals ...
  - Flowering plants
    - \* This opened many new opportunities for *everyone* (insects, mammals, birds), not just the innovators
    - \* **Answer:** Animals evolved to exploit the flowering plants
    - \* **Answer:** Flowering plants and animals *co-evolved*
    - \* **Answer:** Natural defenses and exploits
    - \* **Answer:** Pollination and dispersal

## Hox genes

- **Hox genes** are involved in determining the identity of different body parts
- Taxa with simpler body structures tend to have fewer hox genes
  - Phylogenetic comparisons provide important evidence that hox genes were involved in evolution of complex body plans
- Evidence that new hox genes were largely created by **gene duplication** events
  - A rare but important type of mutation
  - Natural selection must have played a role in the development of this much complexity

## Gene duplication

- One or more genes may be accidentally duplicated so that the genome has two copies of each gene
- This may make the organism less efficient, and thus be selected against
- It may also allow for innovation. Why?
  - **Answer:** Because one copy can continue to do the old function, while the other evolves a new function

## 3.2 Mass extinctions

- Five major mass extinctions so far
- The last one (and maybe others) was caused by a cosmic impact
  - It could happen again
- Are we in the middle of a mass extinction now?
  - Extinction rate very high over the last 400 years
  - Time will tell, but we should be worried

## Conclusion

- The best way to reconstruct evolution is with a *model* of how it occurred
  - Molecular information, and computer modeling has changed our view of the tree of life
  - life is not really a tree
    - \* genetic information can be transferred
    - \* sexual mixing occurs at different scales
- Scientists use many clues to figure out the history of life on earth
  - Fossils, geology, phylogeny
- Life has diversified both gradually, and with dramatic episodes
  - Adaptive radiations, mass extinctions