

# **BIOR93 Module 5**

## **Phylogenetics: An Introduction**

**Jadranka Rota**

**Biological Museum**

**Department of Biology**

**Lund University**

**Etka Yapar**

**Biodiversity Unit**

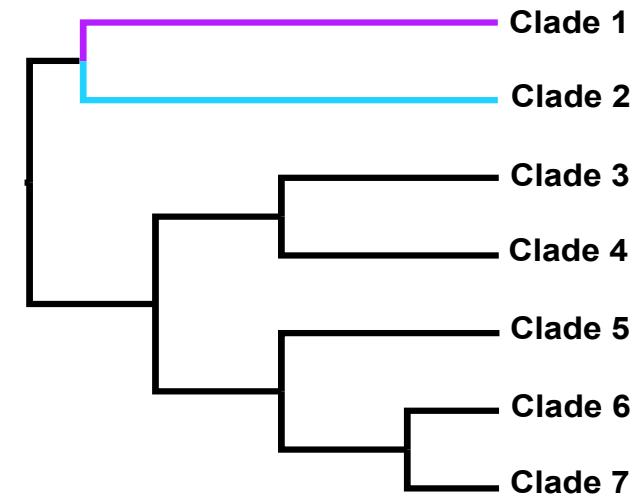
**Department of Biology**

**Lund University**

**Systematic Biology Group**

# Aims

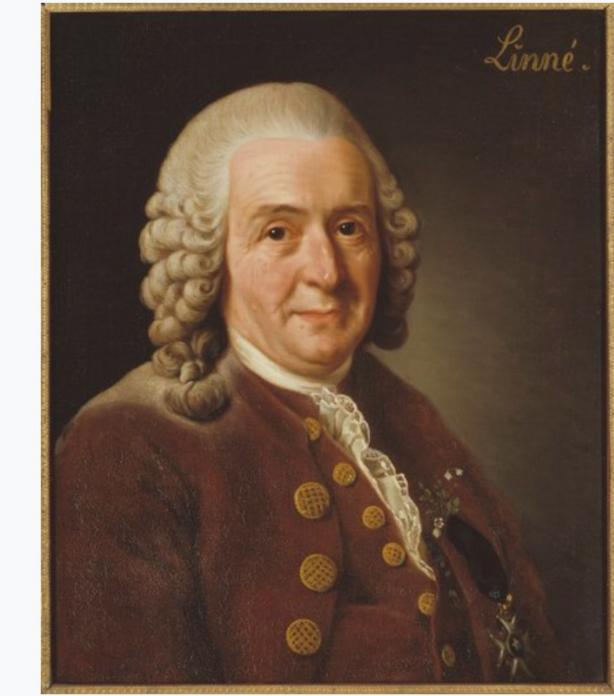
- Cover key concepts in phylogenetics
  - E.g. monophyly, homology, analogy
- Explain why evolutionary history is important in biology
- Understand the basics of statistical phylogenetic inference
- Develop “tree thinking”



# Systematics is ....

- The scientific **study of the kinds and diversity of organisms and of the relationships among them**
- Traditionally: taxonomy (naming and classifying organisms, Greek "taxis" = arrangement, and "nomos" = law)
- Since ca. 1980s: largely based on phylogenetics (first morphological, then molecular)
- Most recently: including phylogenomics
- Provides essential framework for recognition and study of biodiversity and evolution

Carl Linnaeus



Carl von Linné by Alexander Roslin, 1775  
(oil on canvas, Gripsholm Castle)

Known as the “Father of modern taxonomy”  
(Source: Wikipedia)

# Human need for taxonomy

- Naming of organisms around us
  - Makes communication a lot easier!
- Sorting different groups into higher categories
  - Helps us organize the living world around us
  - E.g. porcini (**Karljohanssvamp**) is an edible species of mushrooms, but many other mushrooms are very poisonous, e.g. fly amanita (**Röd flugsvamp**)



Porcini - *Boletus edulis* Bull. (1782)

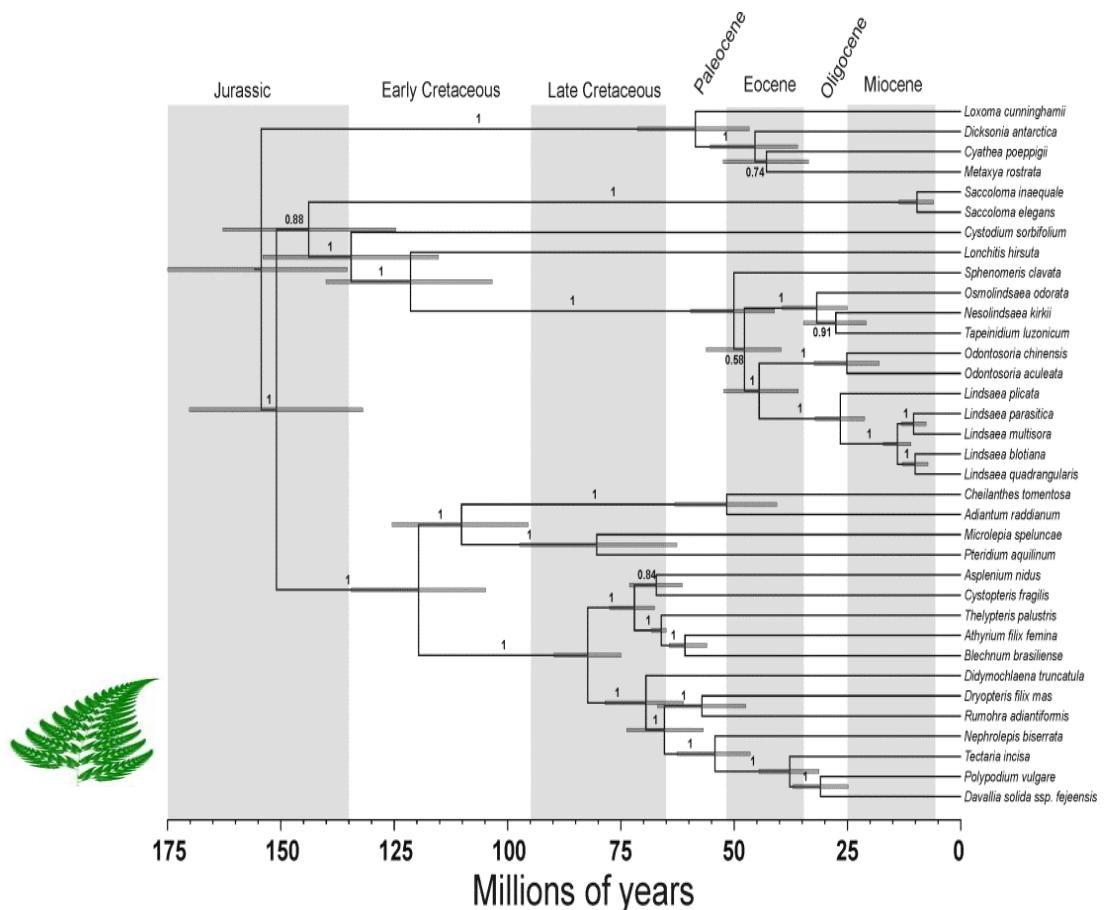


Fly agaric - *Amanita muscaria* (L.) Lam. (1783)

**Note:** one taxon, many taxa (in some Greek words singular ending is –on, plural is –a, e.g. also phenomenon, phenomena)

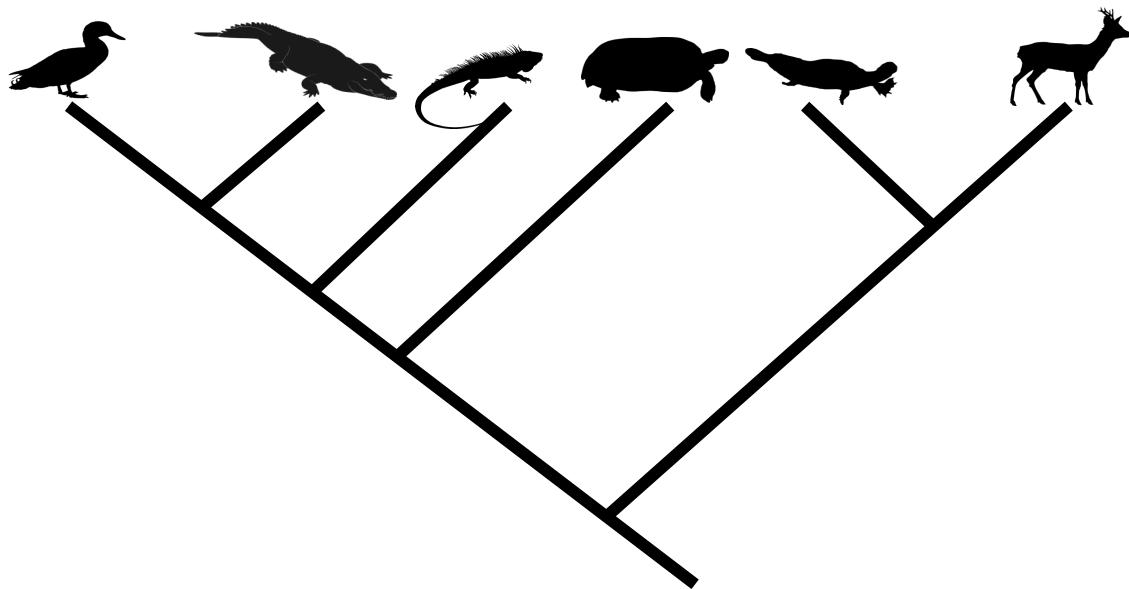
# Systematics also includes...

- The study of character evolution
- The study of molecular evolution
- The study of speciation/extinction dynamics
- The study of historical biogeography
- The study of the temporal framework of evolution



# Evolutionary History

- How do we learn about the evolutionary history of organisms?
- Why should we care about it?

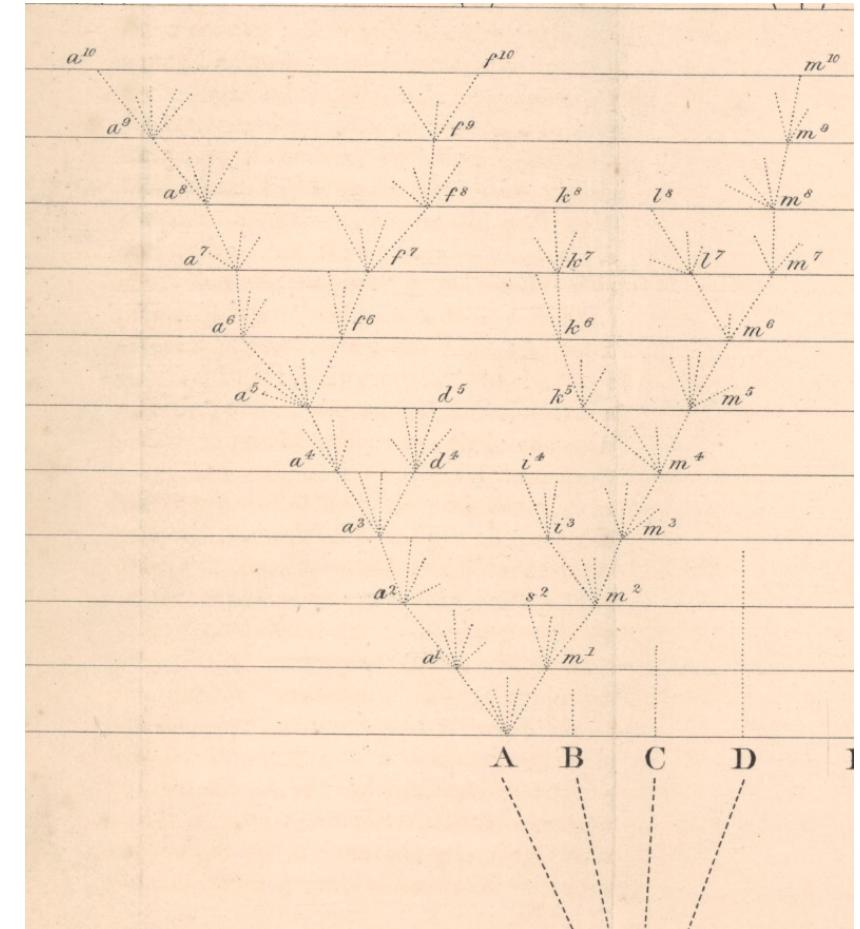


**“Nothing in biology makes sense except in the light of evolution”**

- Theodosius Dobzhansky, essay written in 1973

**“Nothing in evolution makes sense except in the light of phylogeny”**

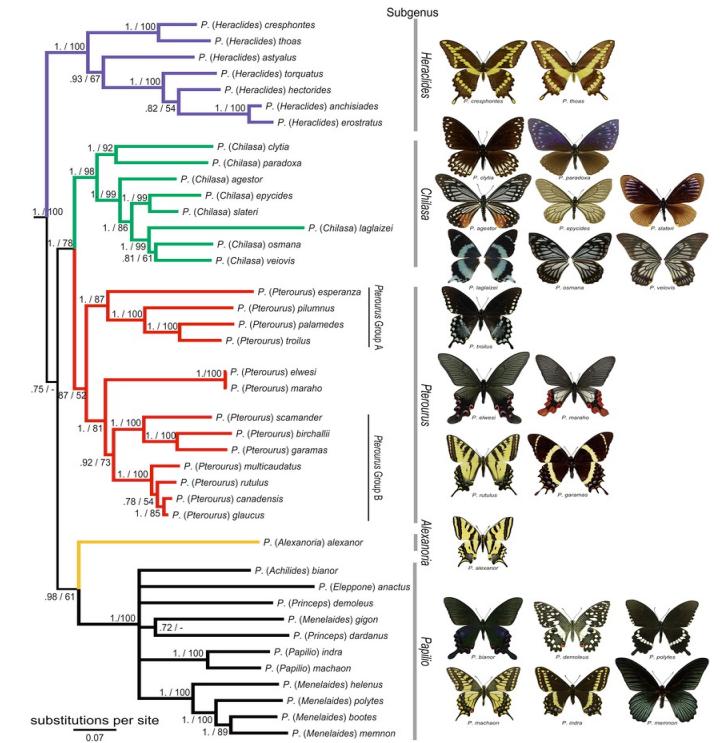
- Quote from Jay Savage, in 1997 Society of Systematic Biology Presidential Address



The only figure in Darwin's On the Origin of Species

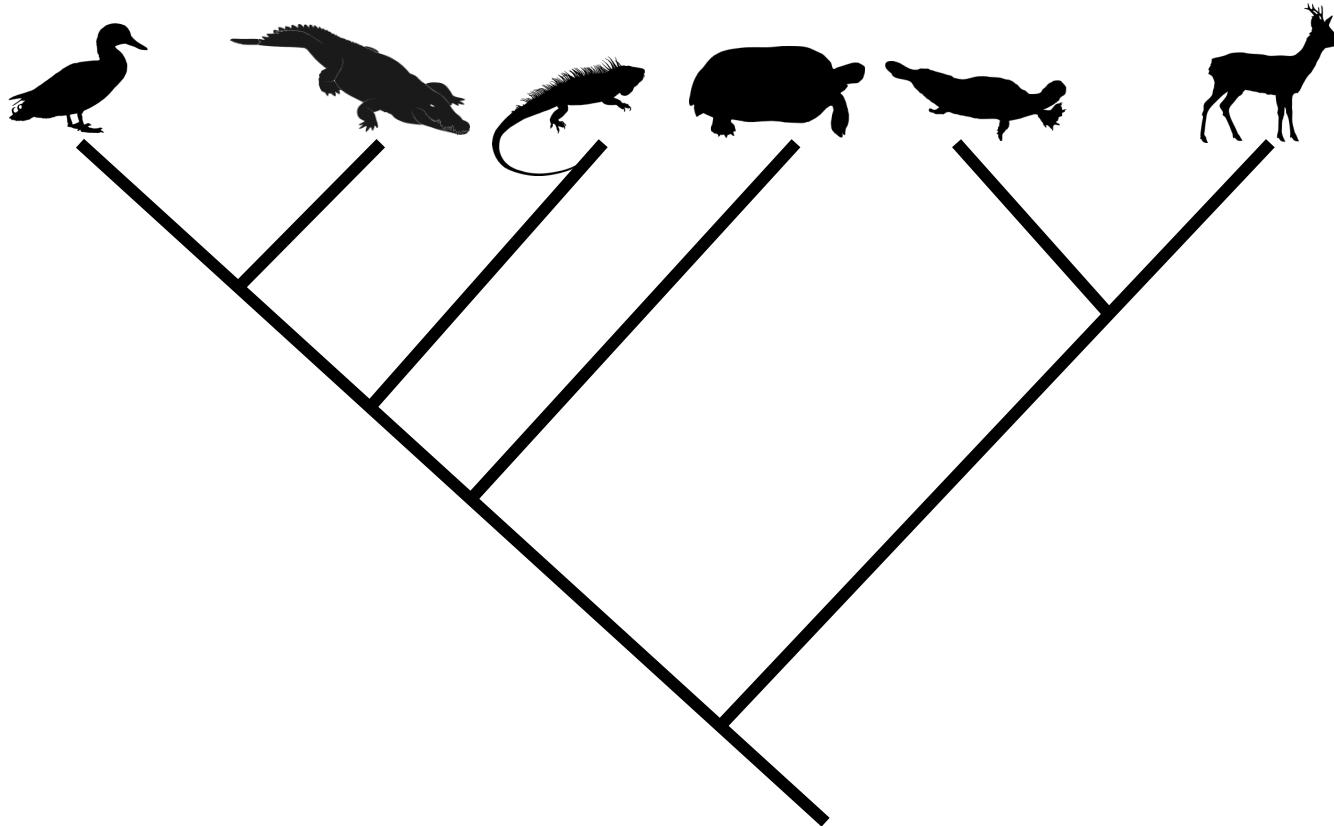
# The very basic facts

- What we see today in nature is the outcome of what has happened in the past
- Ecology and evolution are inseparable
- “Species” or “genes” are not individual entities without any connections to other species or genes
  - phylogeny



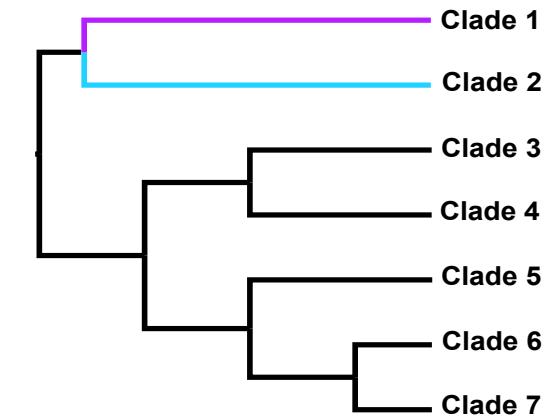
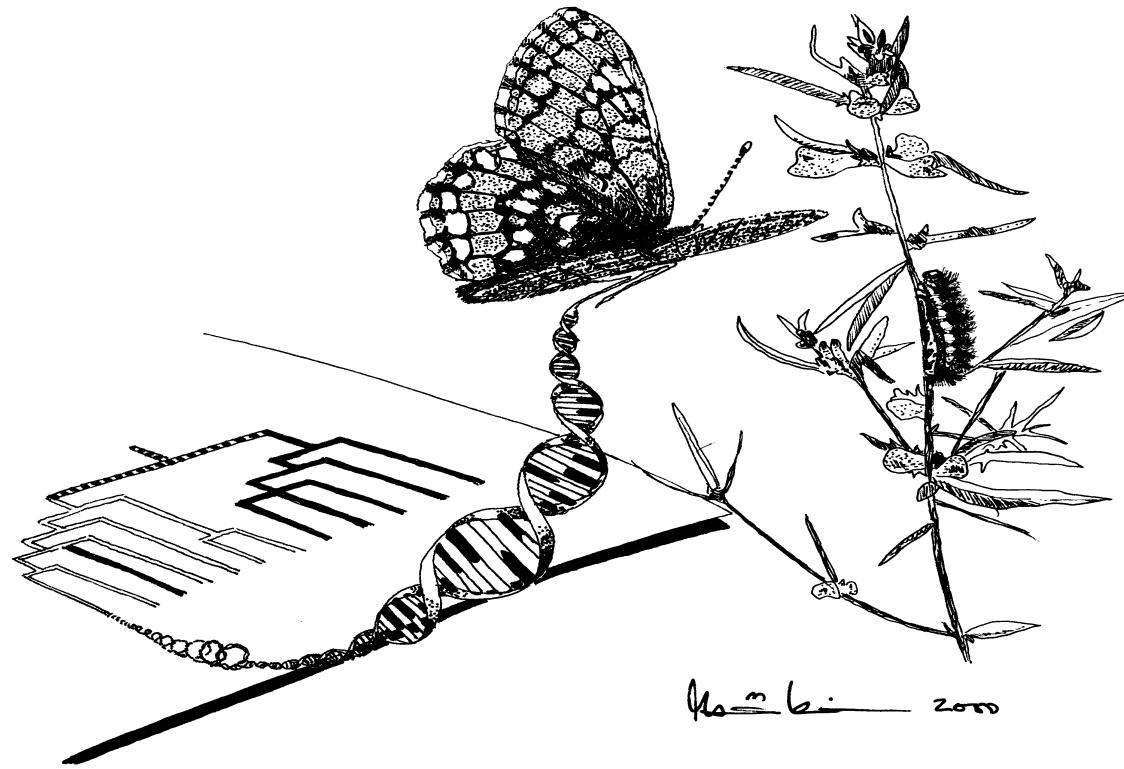
# What is a phylogeny?

- A phylogeny is the historical genealogy of a group of species



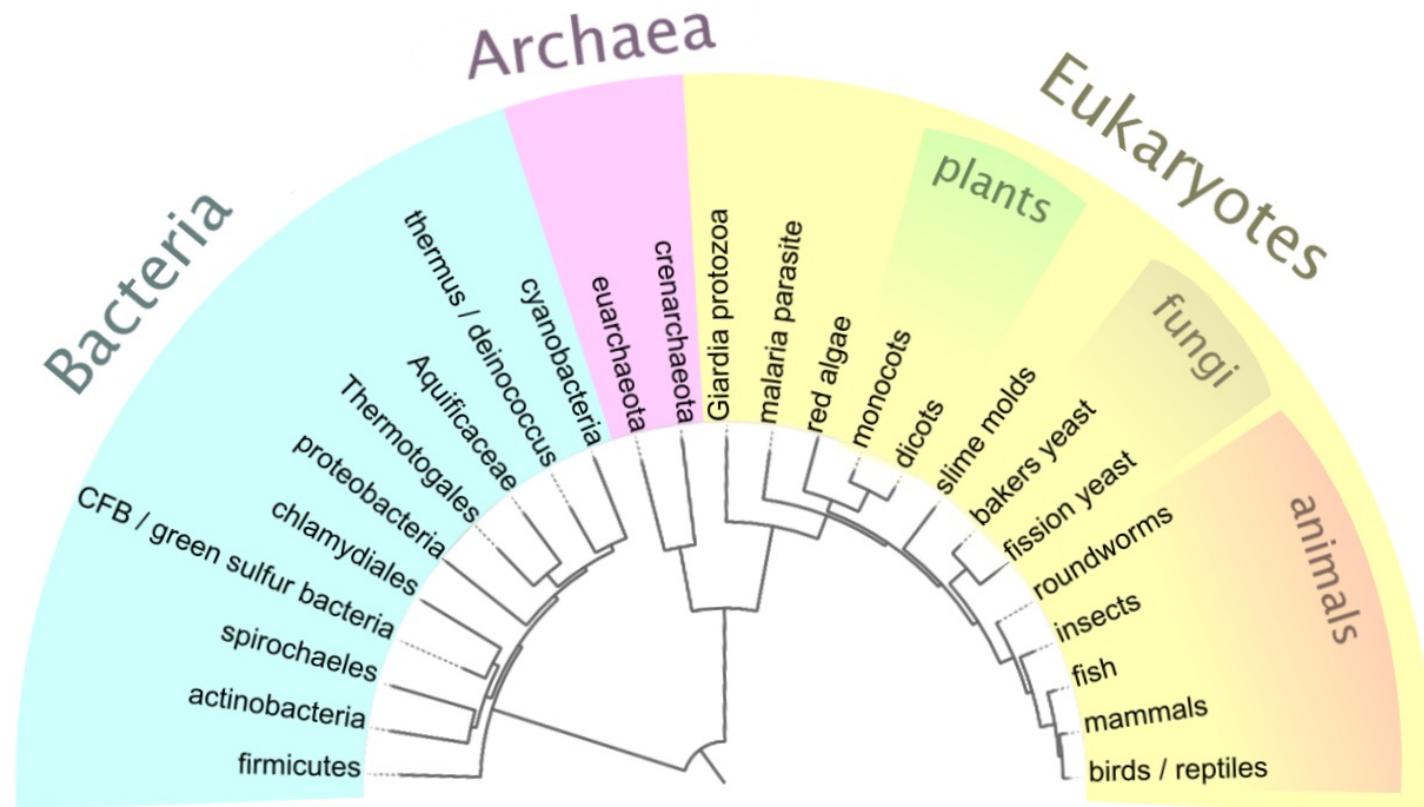
# A phylogeny is an inference

- Envisioned as a dichotomously branching tree
- A phylogeny cannot be observed
- A phylogenetic hypothesis can be inferred from observed data



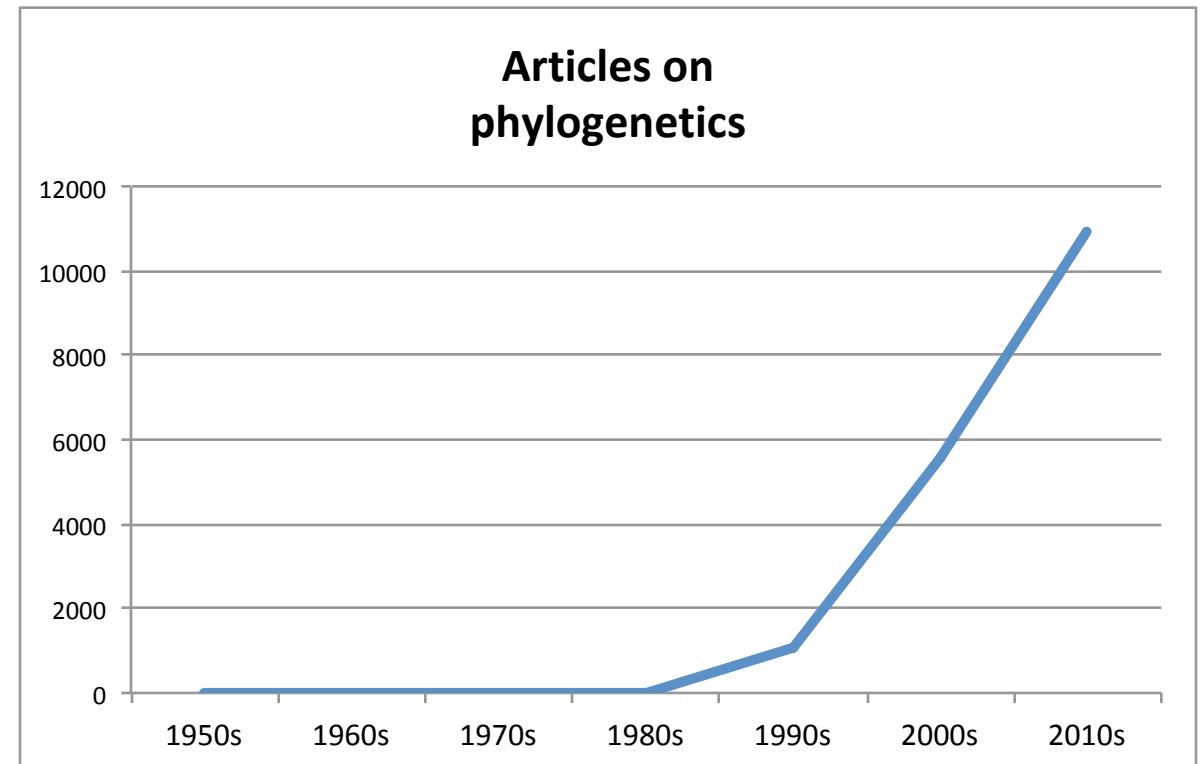
# What we are after

- **Phylogenies – the Tree of Life**
- **With phylogenies we are attempting to get a good working framework for Life**
- **Getting to the root of how evolution has worked**



# The rise of systematics

- Within the last 25 years the number of phylogenetic studies has skyrocketed

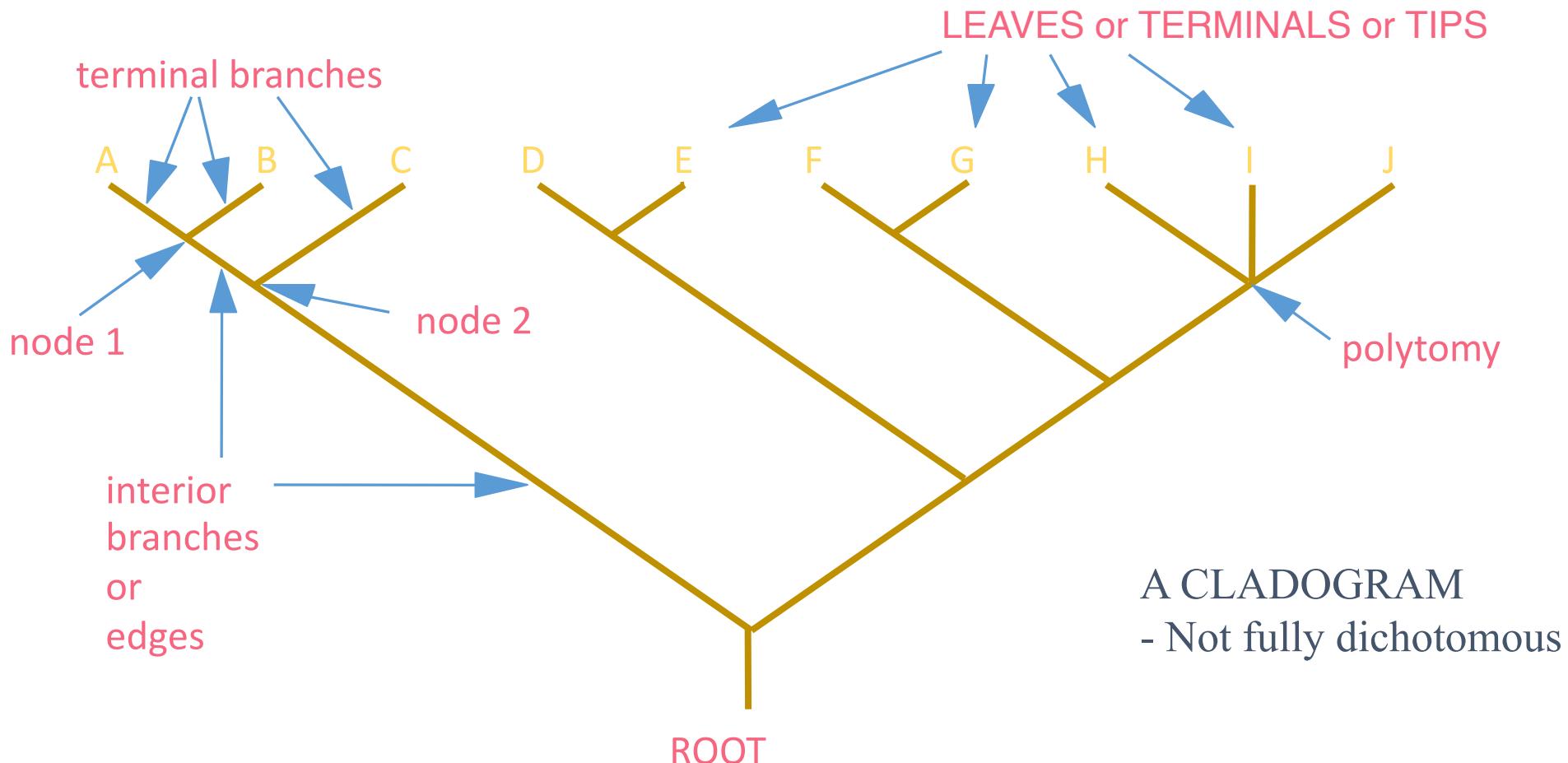


Web of Science, topic = phylogenetics, search done on Aug 29, 2022

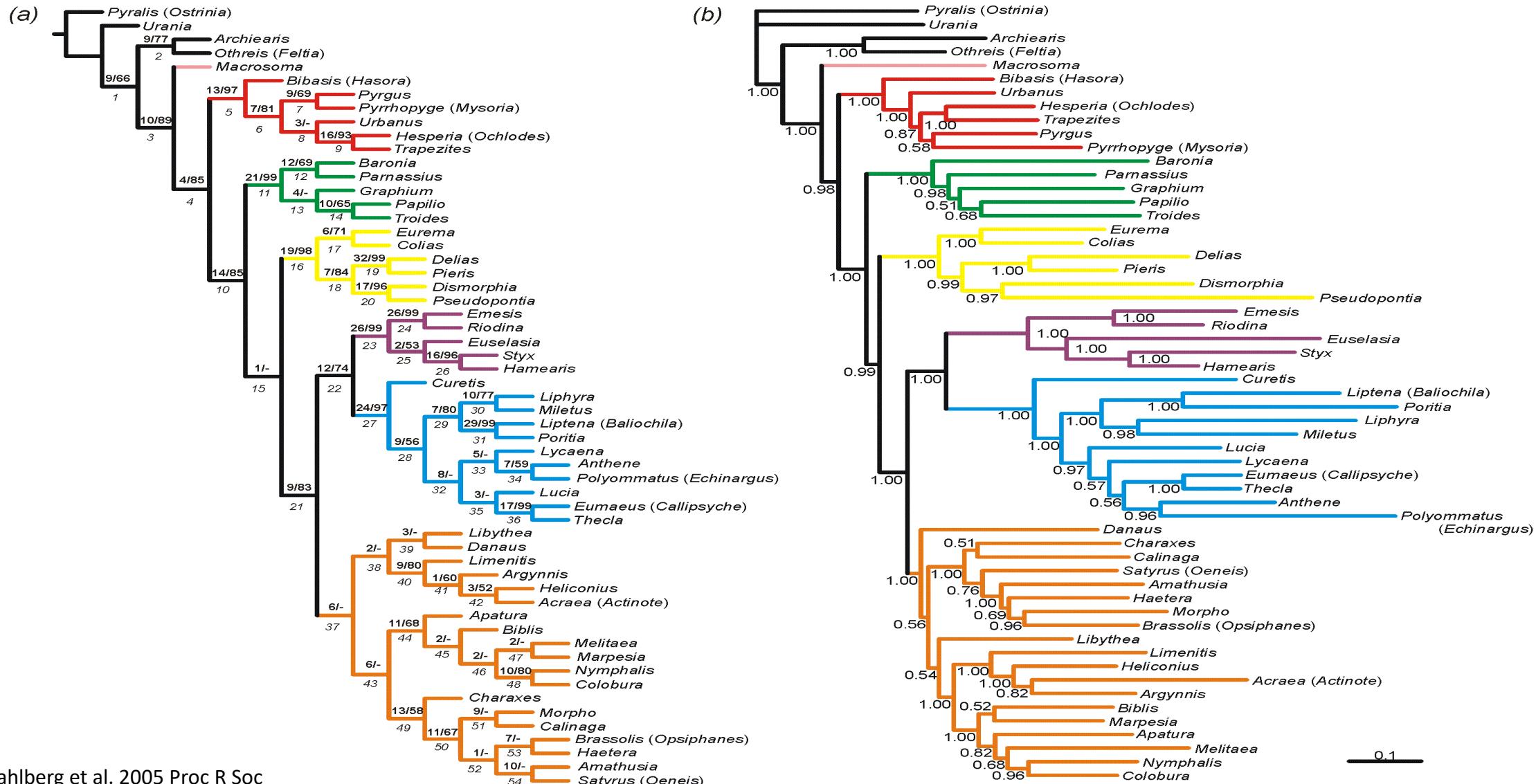
# Some basic concepts

- **Cladogram** – a tree diagram which depicts a hypothesised evolutionary history (topology)
- **Phylogram** – a tree which indicates by branch length the degree of change believed to have occurred along each lineage (topology with informative branch lengths)
- **Chronogram** – a tree in which branch lengths are directly in proportion to time (a type of an ultrametric tree – all tips are equidistant from the root)

# Phylogenetic Trees

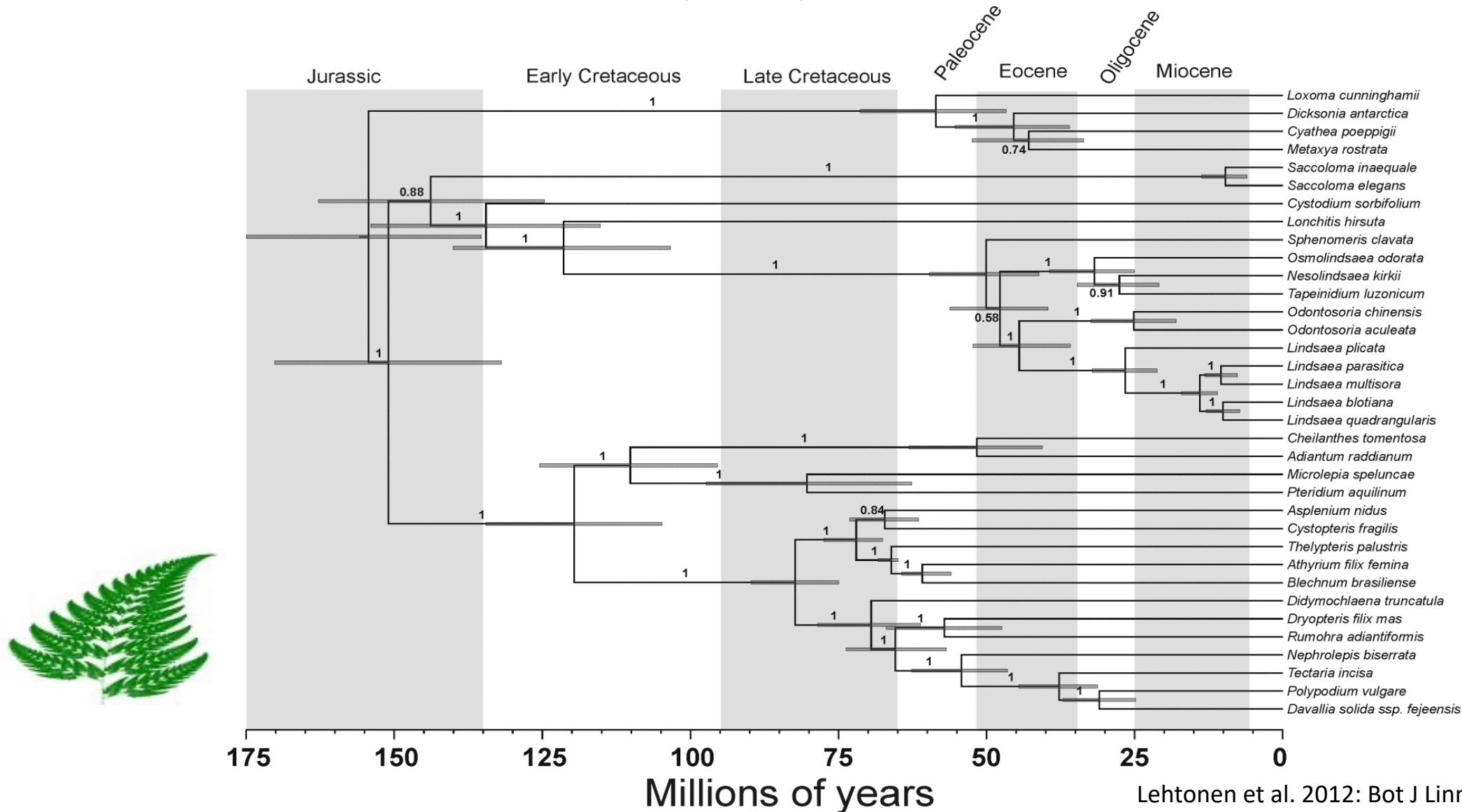


# Cladograms and phylogenograms

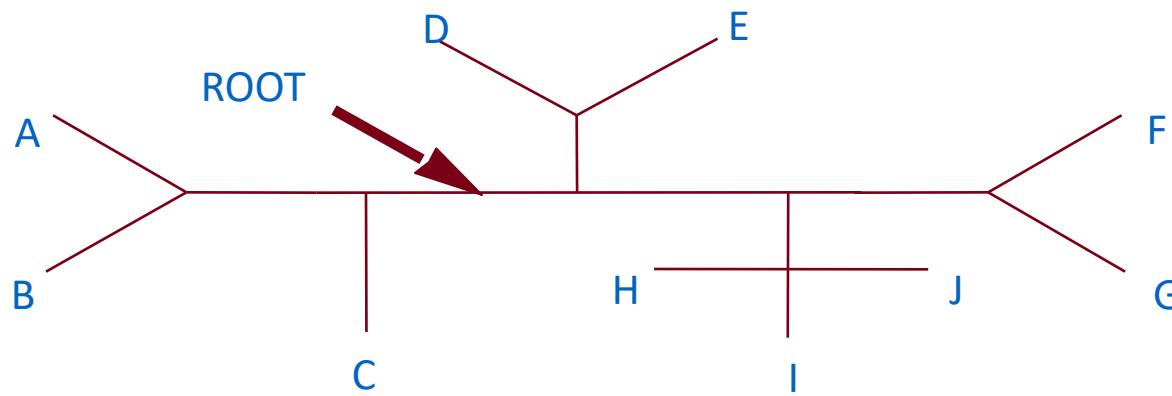
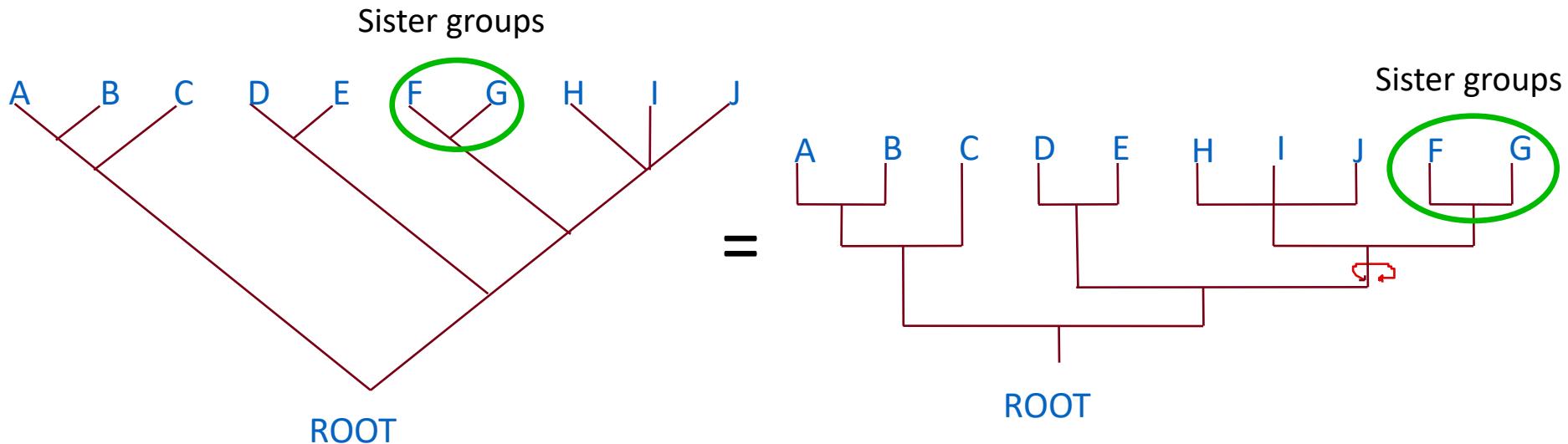


# Chronogram

(an ultrametric tree – all tips equidistant from the root)



# Trees - Rooted and Unrooted



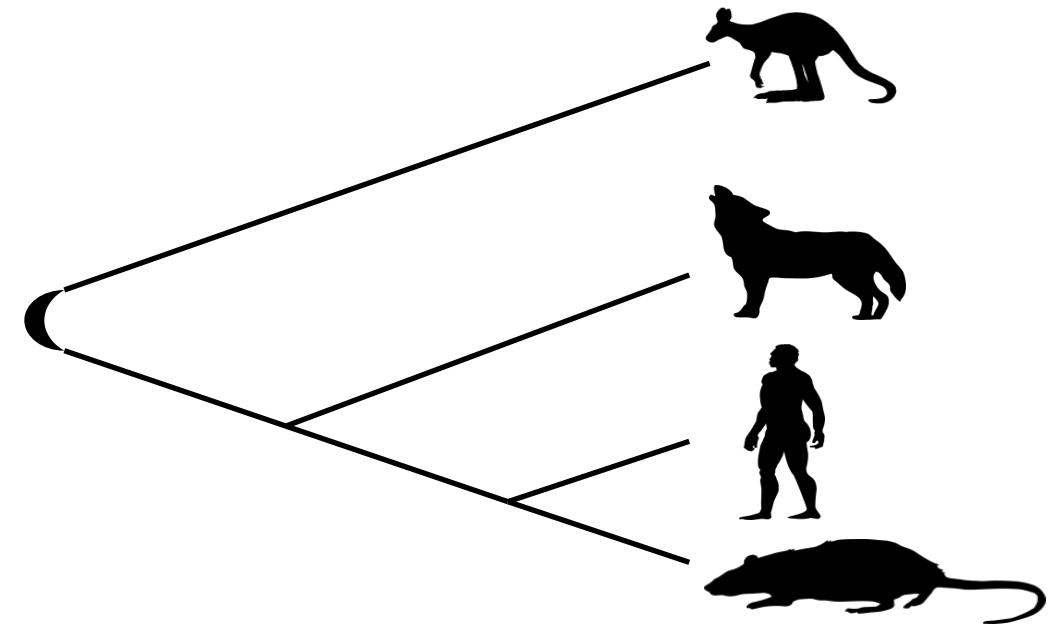
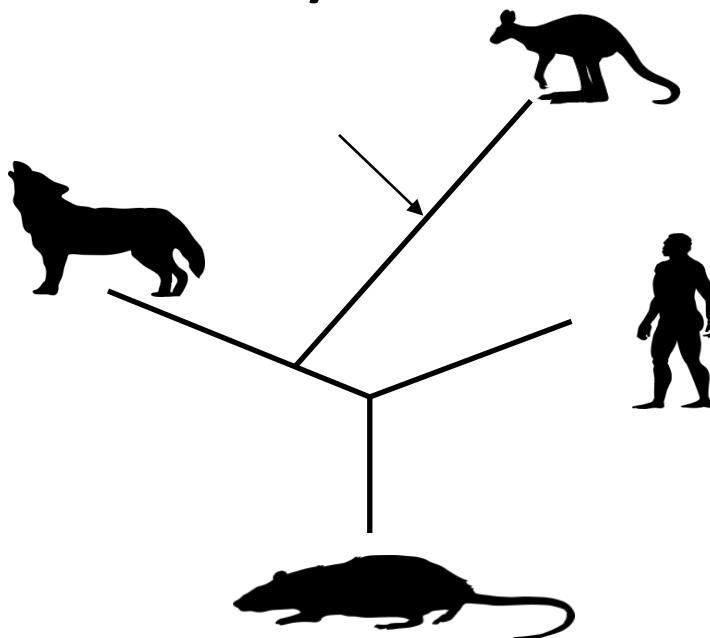
# Rooting a tree

- **Rooting a tree using outgroups**
  - Commonly we include several outgroups
  - Place the root on the branch leading to the outgroup taxon
- **Other ways of rooting a tree**
  - Assume a molecular clock
  - Midpoint rooting (root on the longest branch)



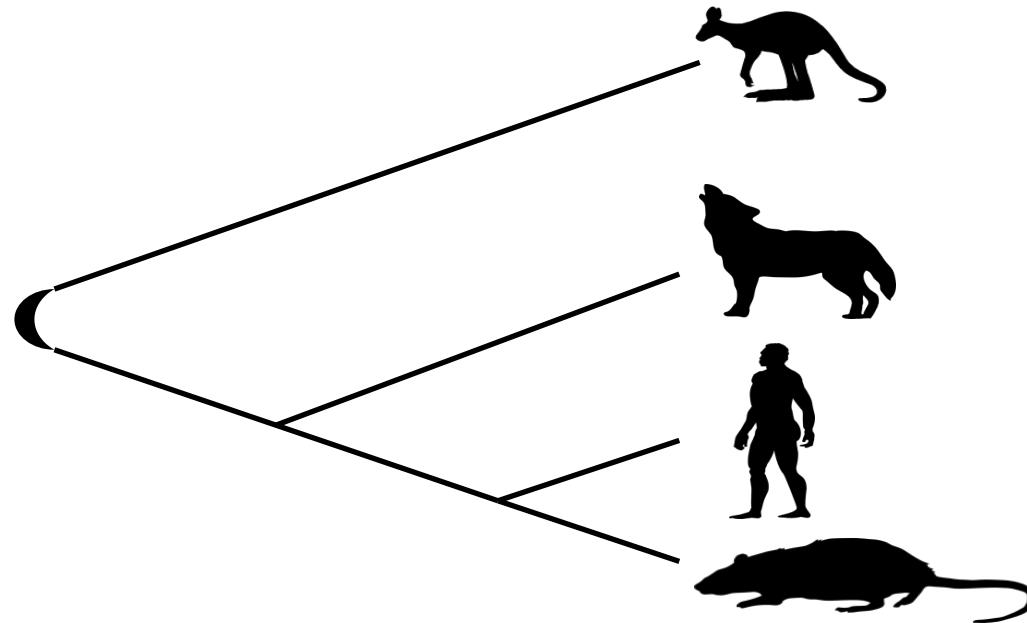
# Outgroup rooting of unrooted trees

- Outgroup – related group that definitely diverged earlier (palaeontological evidence)
- Not too distantly related (tree method becomes unreliable if it is too distant)

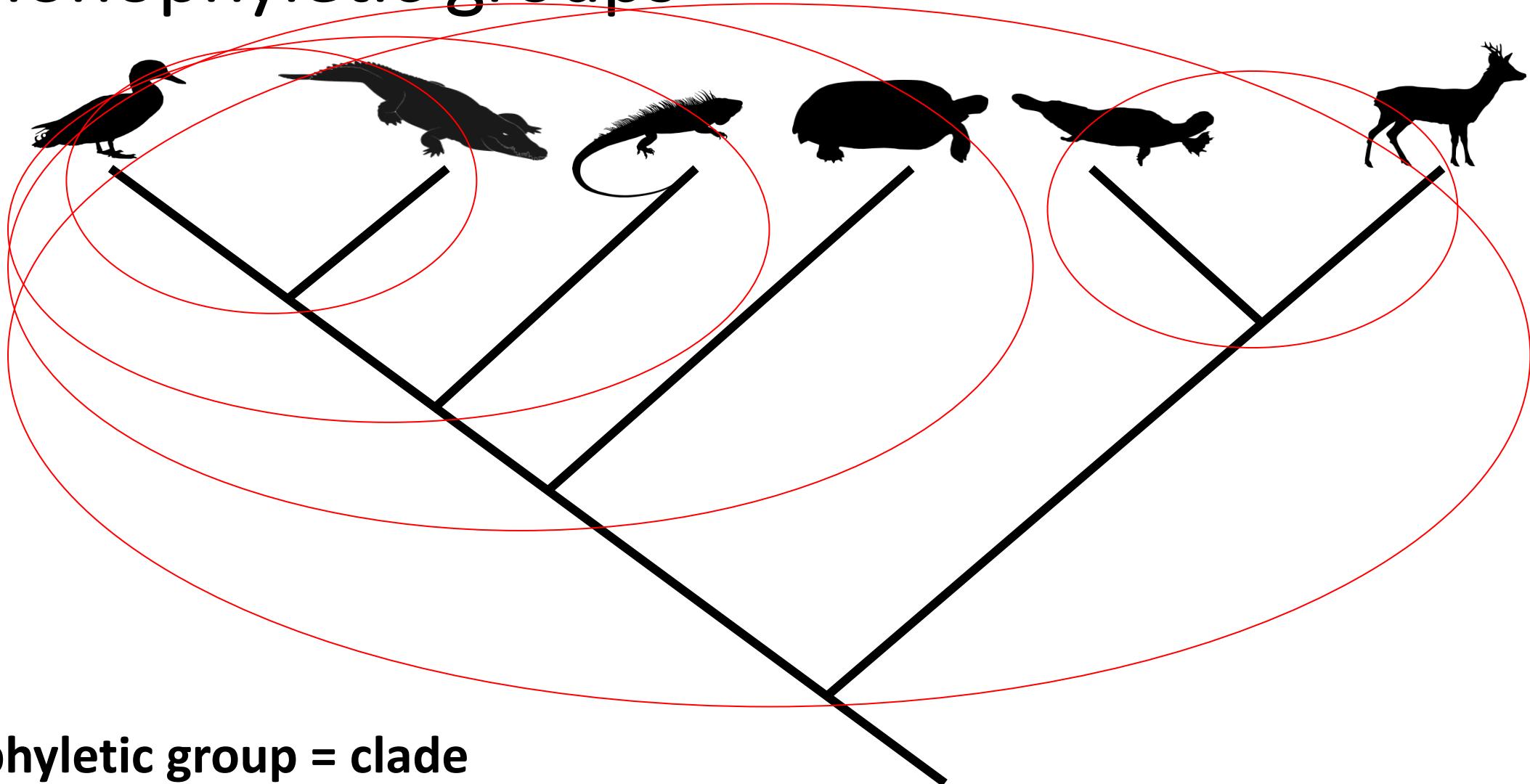


# Phylogenetic systematics

- Uses tree diagrams to portray relationships based upon recency of common ancestry
- **Monophyletic groups (clades)** – contain species which are more closely related to each other than to any outside of the group, including the MRCA (most recent common ancestor)



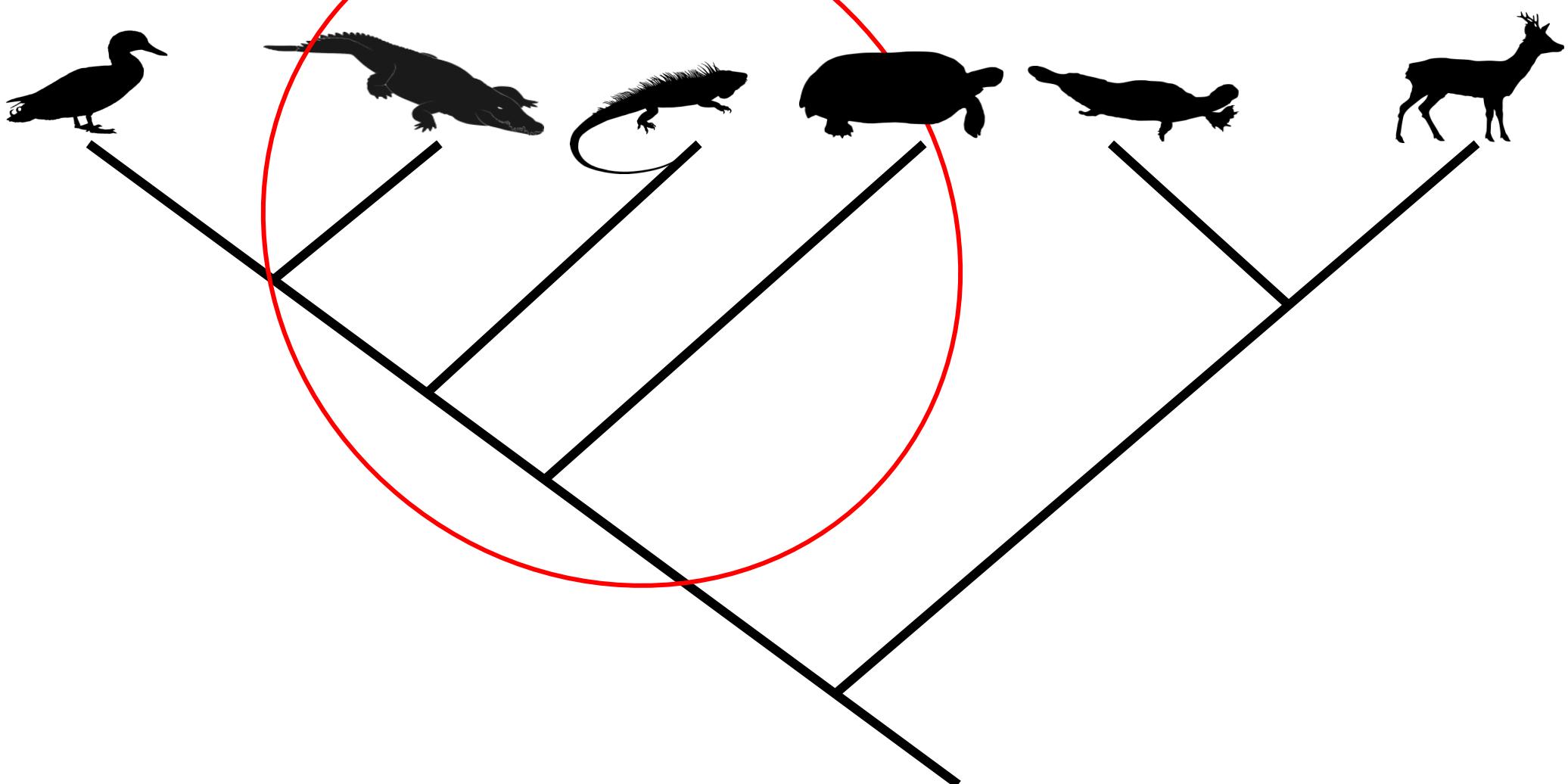
# Monophyletic groups



**Monophyletic group = clade**

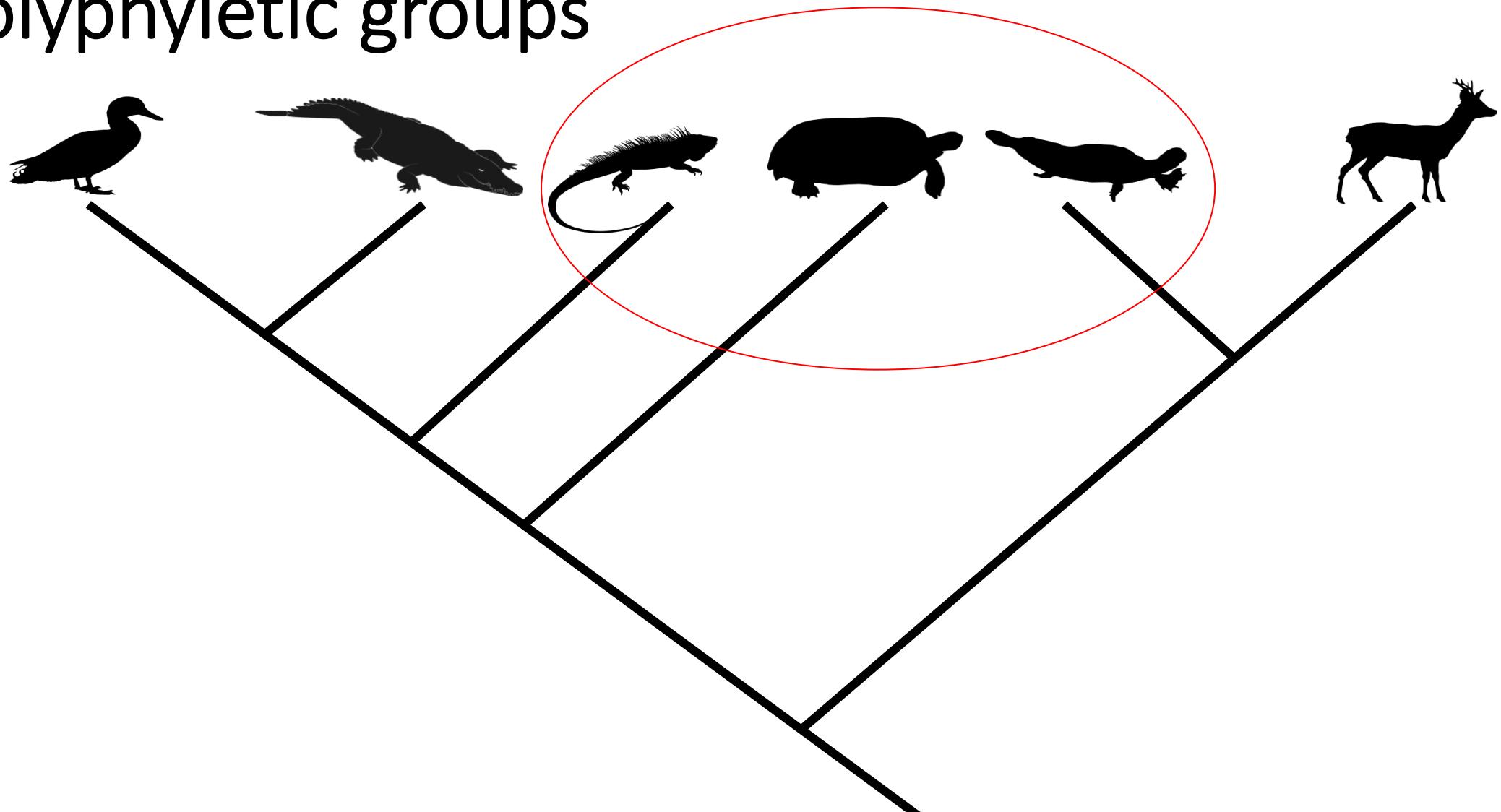
- Include an ancestor and all of its descendants

# Paraphyletic groups



- Include ancestor and some but not all of its descendants

# Polyphyletic groups



- Include some but not all of the descendants and exclude the ancestor

# Sister groups

- By definition, sister groups are of equal age
- Common mistake
  - The sister group that has fewer species is referred to as basal
  - Possible to have nodes that are more basal than other nodes, but not lineages compared to their sister group
  - Rather than saying “this group is basal”, one should say “this group is sister to all other lineages”

This question is actually not yet considered resolved!

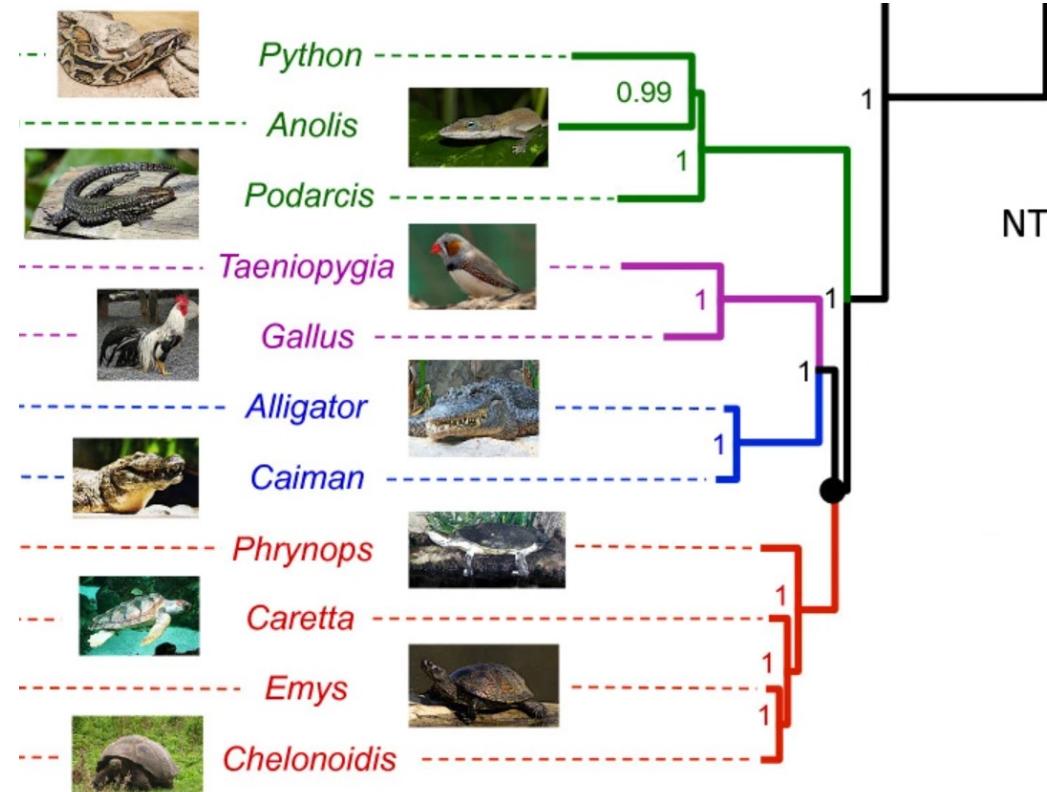
Research article | Open Access | Published: 27 July 2012

**Phylogenomic analyses support the position of turtles as the sister group of birds and crocodiles (Archosauria)**

Ylenia Chiari ✉, Vincent Cahais, Nicolas Galtier & Frédéric Delsuc ✉

*BMC Biology* 10, Article number: 65 (2012) | [Cite this article](#)

32k Accesses | 241 Citations | 50 Altmetric | [Metrics](#)

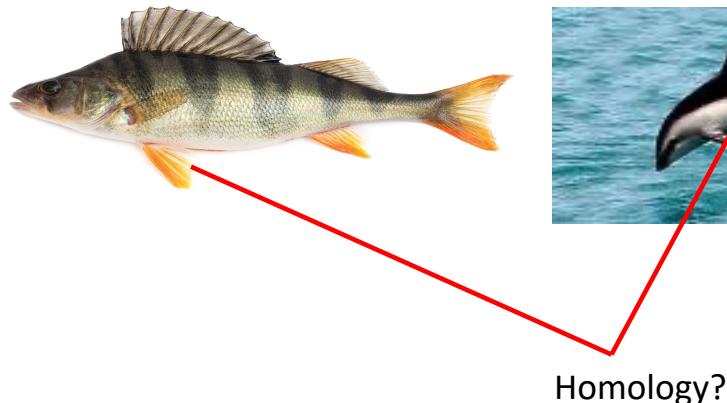


# Some premises underlying phylogenetic inferences

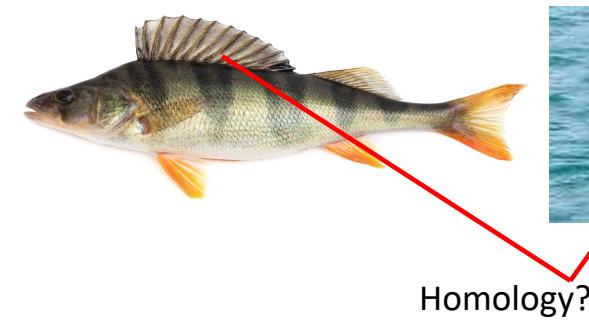
- Phylogenetic inferences are premised on
  - the inheritance of ancestral characters
  - the existence of a shared evolutionary history
- Homology considered as evidence of common ancestry
- A tree-like model of evolution
  - There are evolutionary processes that don't fit this model, e.g. lateral transfer

# Homology

- The most fundamental concept in inferring phylogeny is **homology**
- We need to be sure the characters we are studying are homologous, i.e. "the same" character in different organisms
- Otherwise our analyses will be misled



vs.



# Owen's definition of homology

**Homologue:** the same organ under every variety of form and function (true or essential correspondence)

Richard Owen 1843

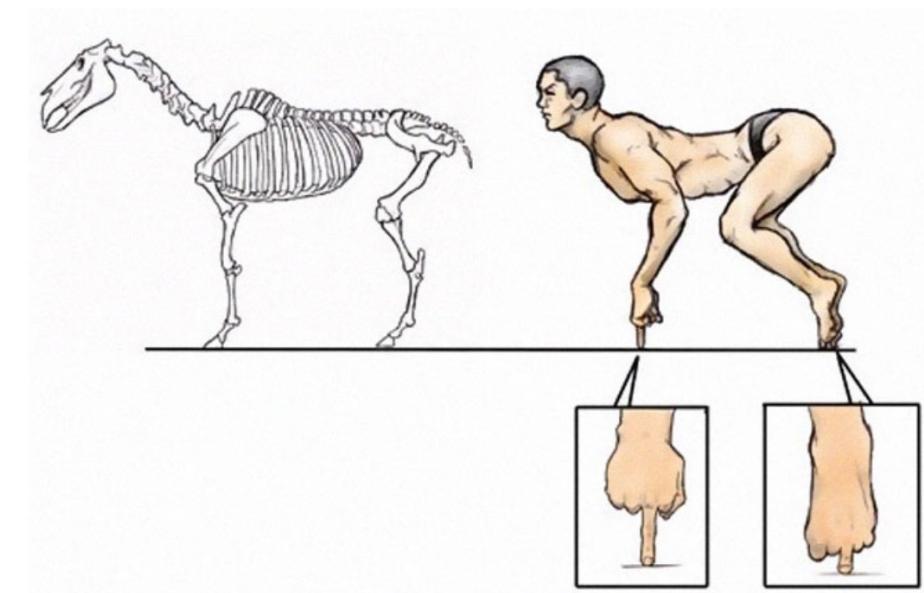
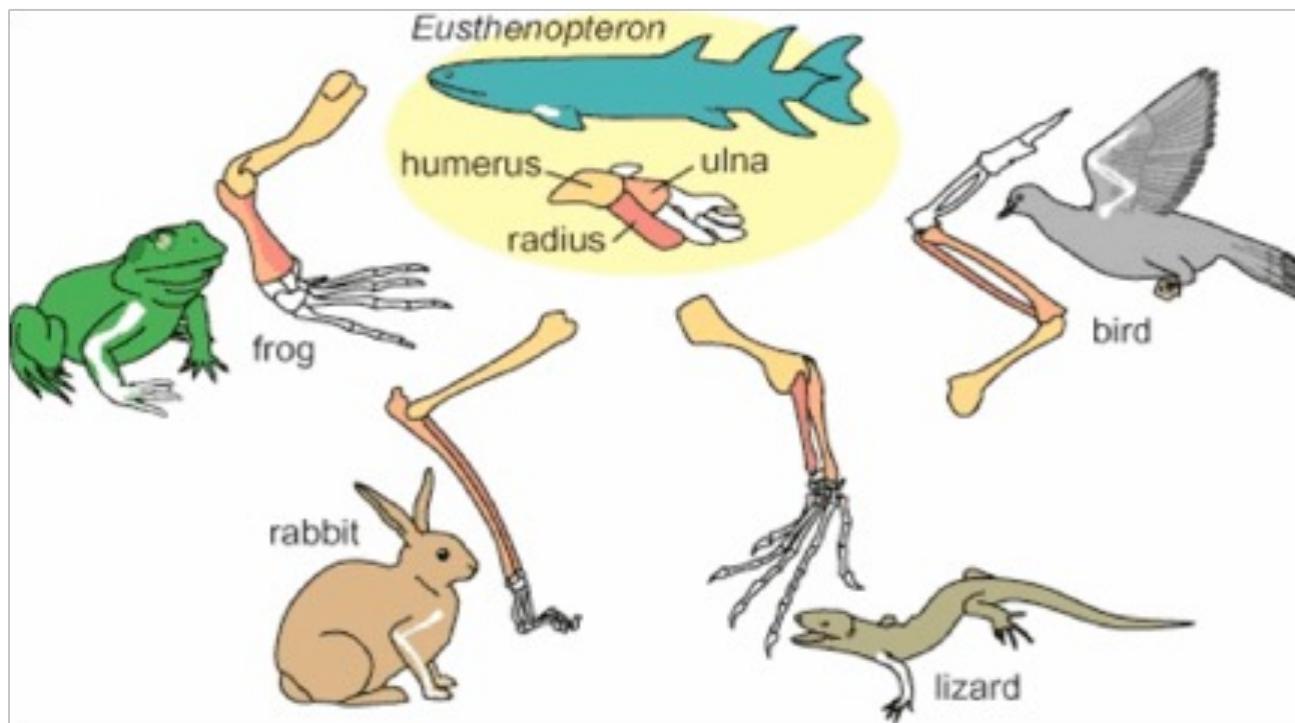


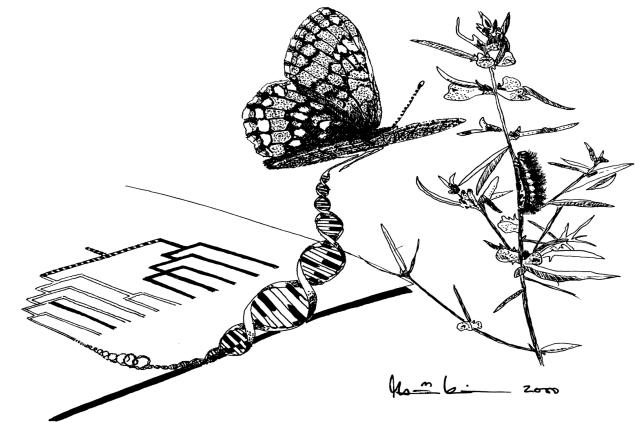
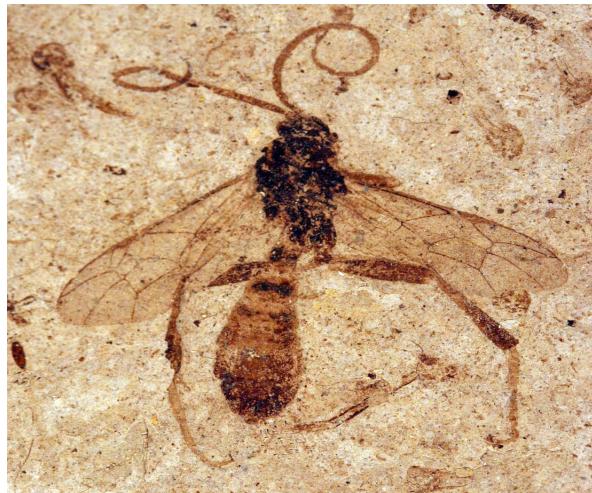
Image source: Satoshi Kawasaki

# Main kinds of data in phylogenetic inference

- Morphological
  - Traditionally used in phylogenetic inference
  - Still necessary for fossils and when molecular data are lacking
  - Can also help when molecular data are ambiguous
- Molecular
  - Most commonly used nowadays
  - Ease of sequencing led to a revolution in molecular phylogenetics

# Phylogenetic analysis is an attempt to infer the past

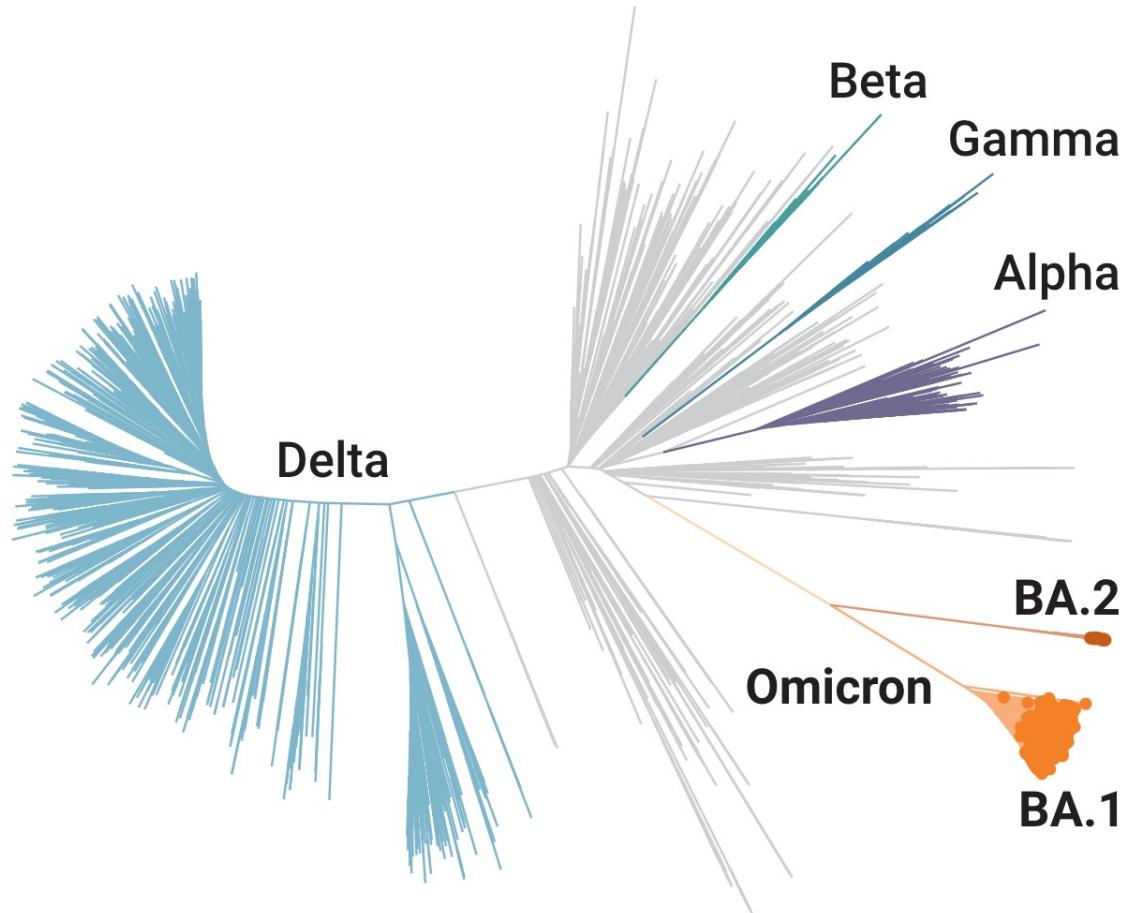
- Inferring a phylogeny is an attempt to produce a best estimate of an evolutionary history based upon incomplete information
- Our direct information about the past is limited
  - Fossil record very incomplete
  - Access to contemporary species and molecules



# Evolutionary History and Its Importance in Biology

# Why do we need phylogenies?

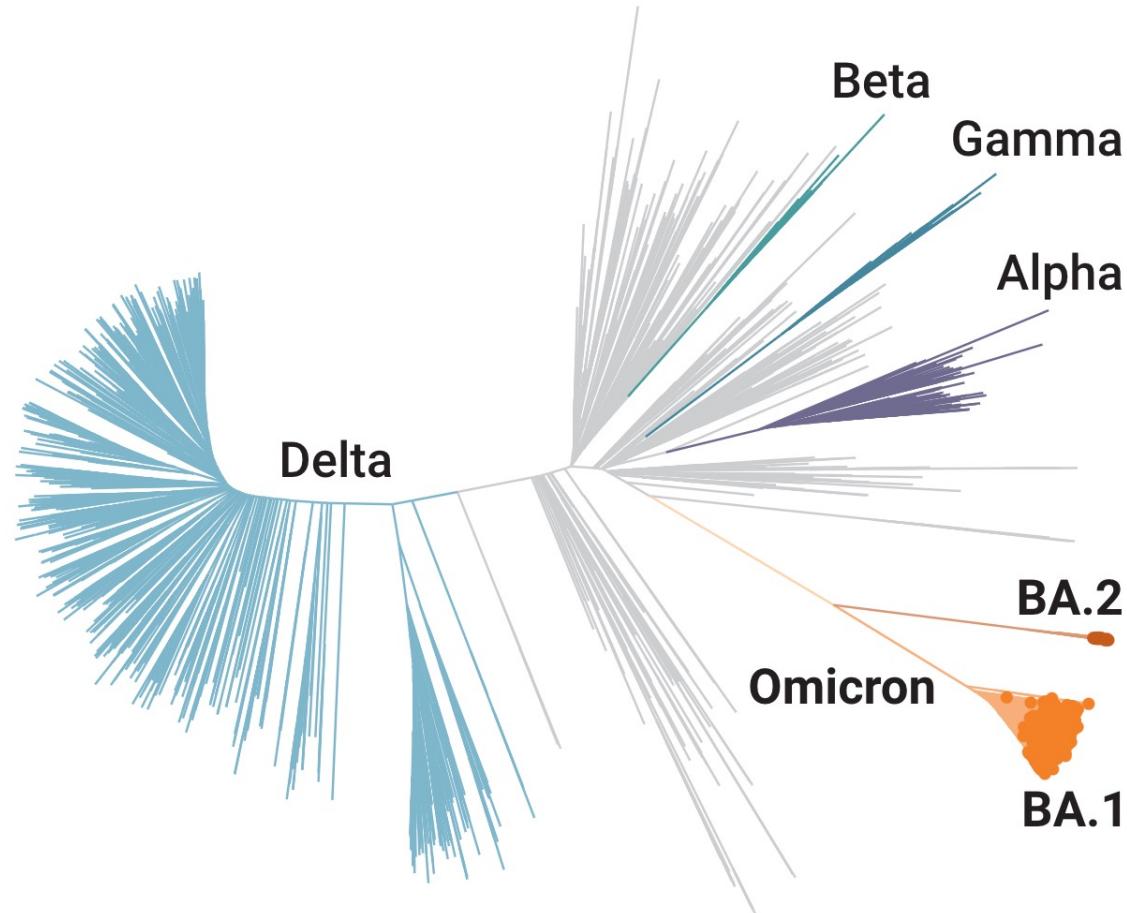
- What is shown here?



# Why do we need phylogenies?

- What is shown here?

Phylogeny of covid-19 strains



# Examples relevant to systematics and evolutionary history

- Snakes and lizards
- Dinosaurs and birds
- Relationships among the three domains of life

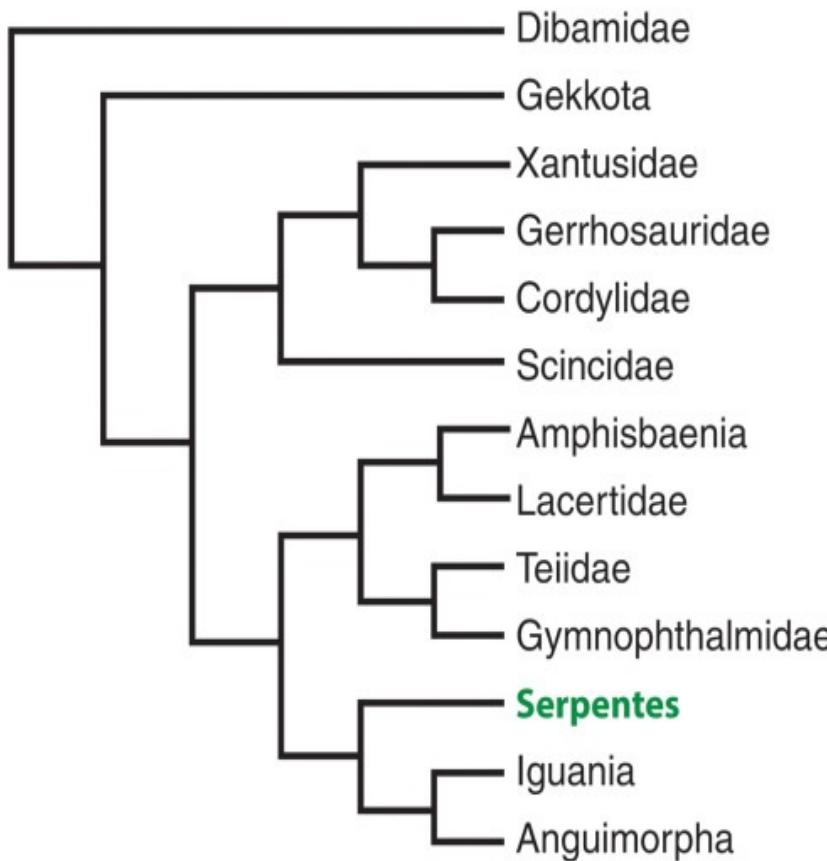
# Which statement is true?

- a) Snakes are the sister group of lizards.
- b) Some lizards are more closely related to the snakes than to other lizards.
- c) Snakes are reptiles that are not closely related to lizards.



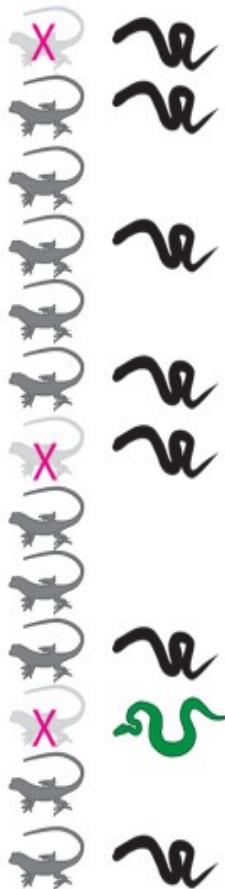
# Snakes are one of several lizard lineages that have lost limbs

a



b

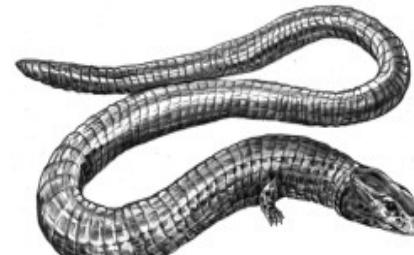
Gallery of snake-like lizards



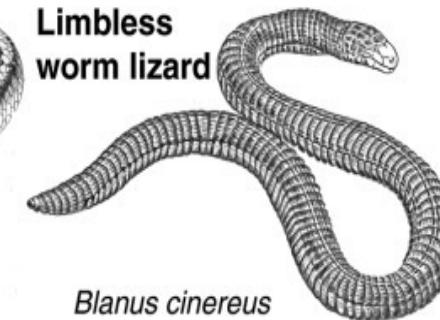
Rear-legged skink



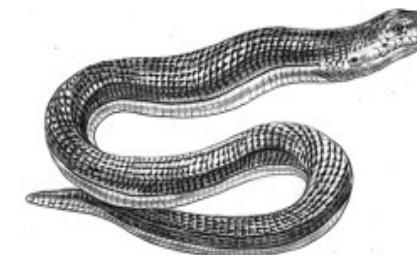
Front-legged microteiid



Limbless worm lizard



Limbless glass lizard



**Snake-like  
body plan**

# Which statement is true?

- a) Snakes are the sister group of lizards.
- b) Some lizards are more closely related to the snakes than to other lizards.
- c) Snakes are reptiles that are not closely related to lizards.



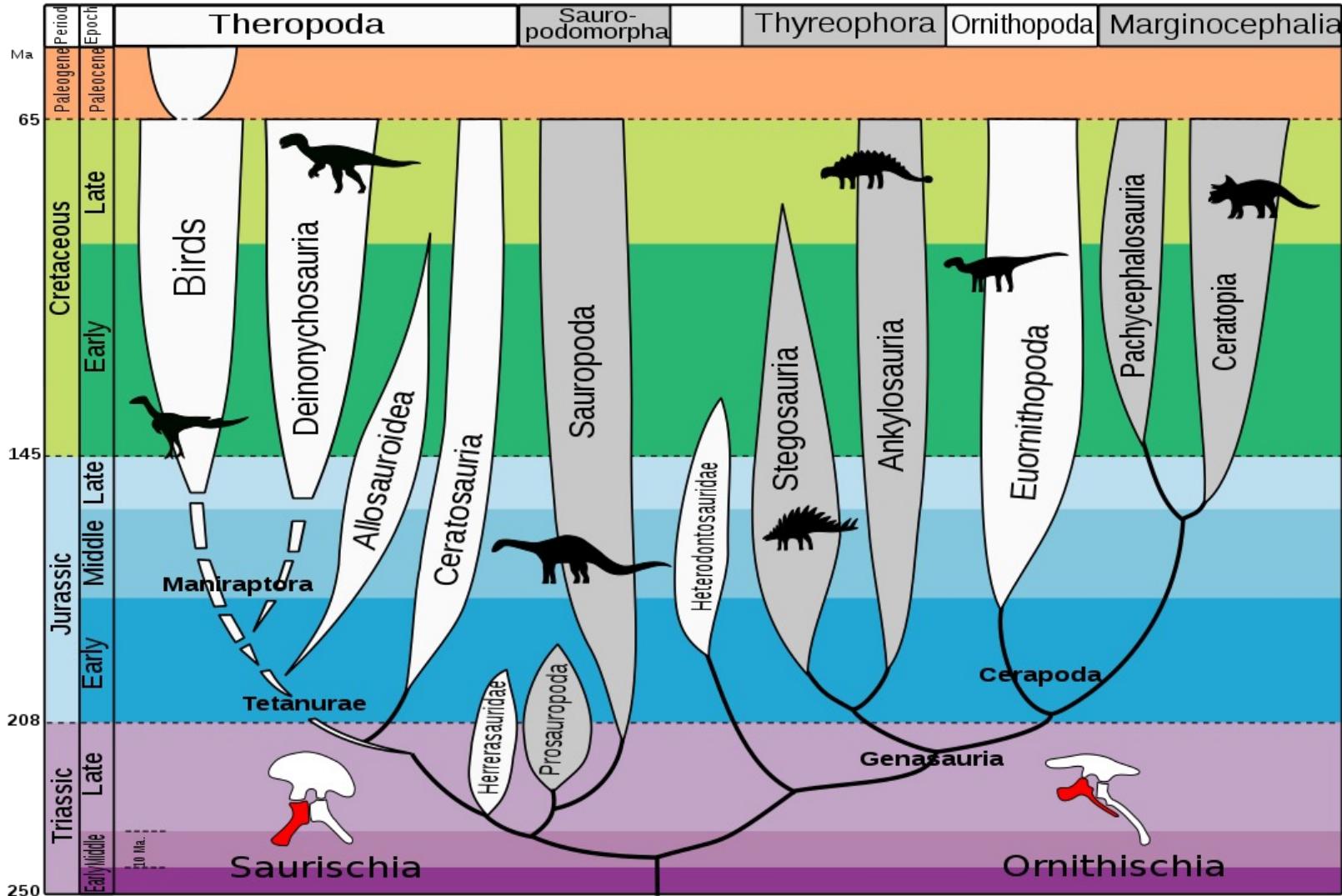
# Dinosaurs...

- a) ... belong to reptiles but are not closely related to birds.
- b) ... are the sister group of birds.
- c) ... are paraphyletic with respect to birds.



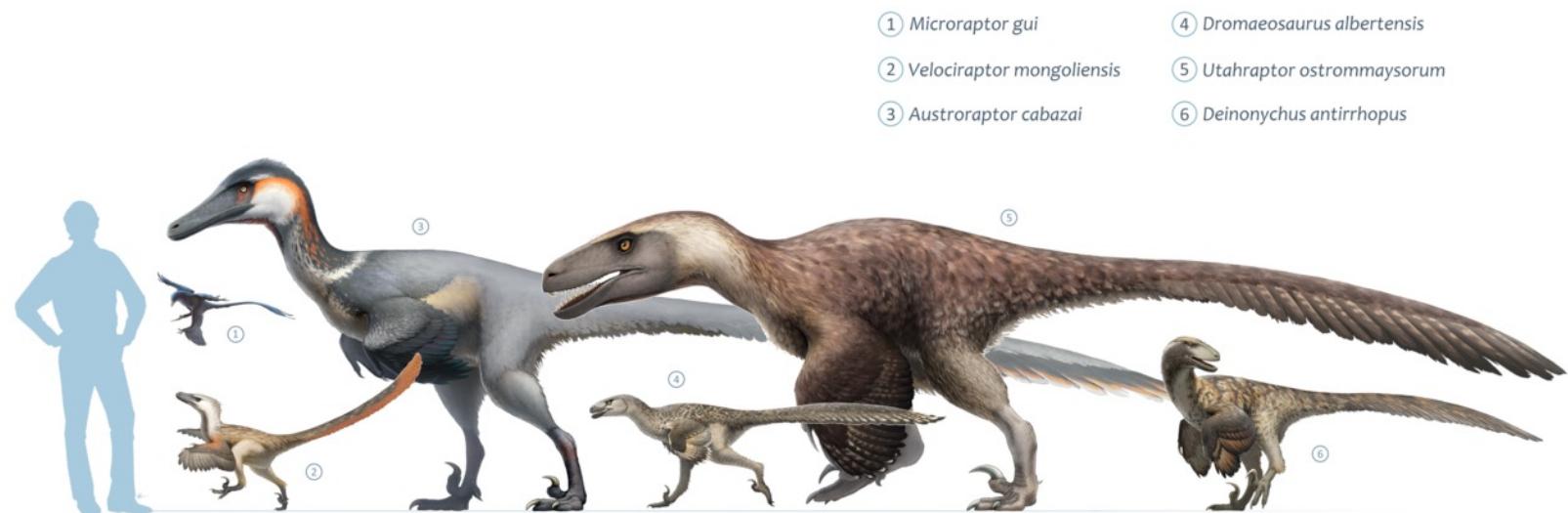
# Are dinosaurs extinct?

- Avian dinosaurs
- Non-avian dinosaurs
- Dinosaurs are paraphyletic with respect to birds
- Pseudoextinction

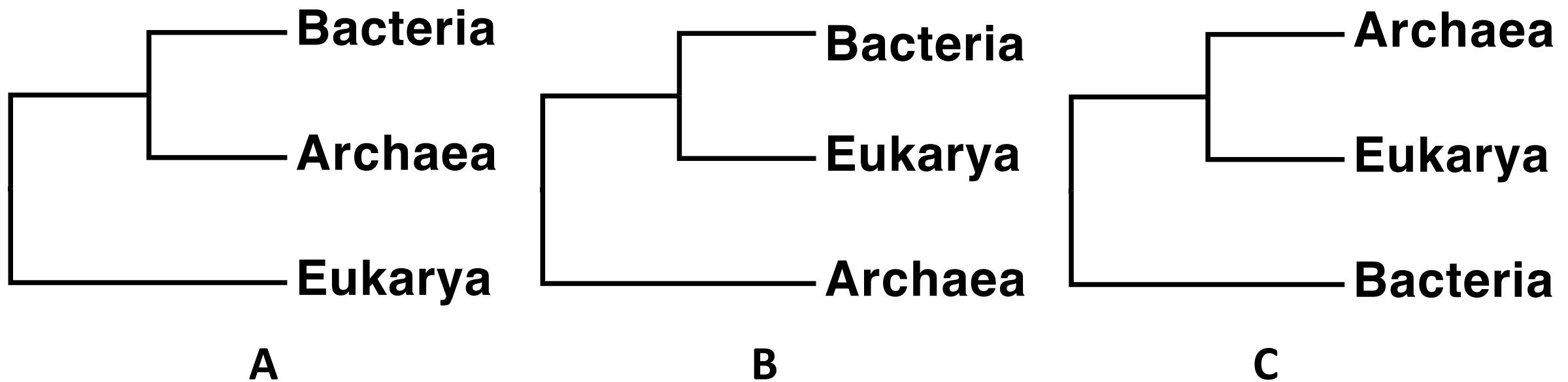


# Dinosaurs...

- a) ... belong to reptiles but are not closely related to birds.
- b) ... are the sister group of birds.
- c) ... are paraphyletic with respect to birds.

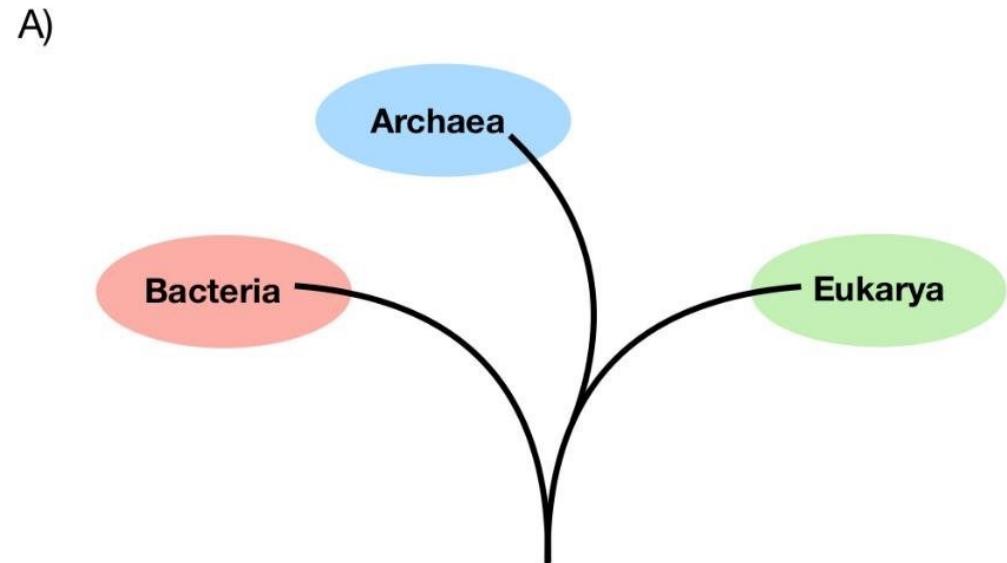


Best hypothesis for the relationships among the three domains of life is:

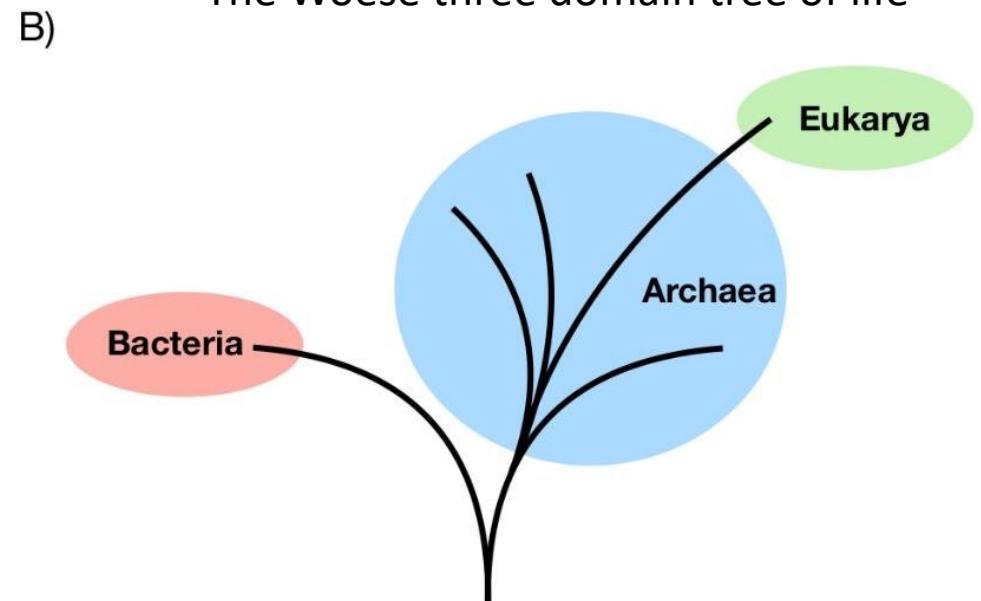


# Origin of Eukarya

- **Eukaryotes**
  - Organisms whose cells have a nucleus enclosed within a nuclear envelope
  - Plants, animals, fungi + a lot of single-cell lineages
- **Which lineage is the sister lineage to Eukarya?**
- **Why is this interesting?**
  - Where does the complex cell of eukaryotes originate from?



The Woese three domain tree of life

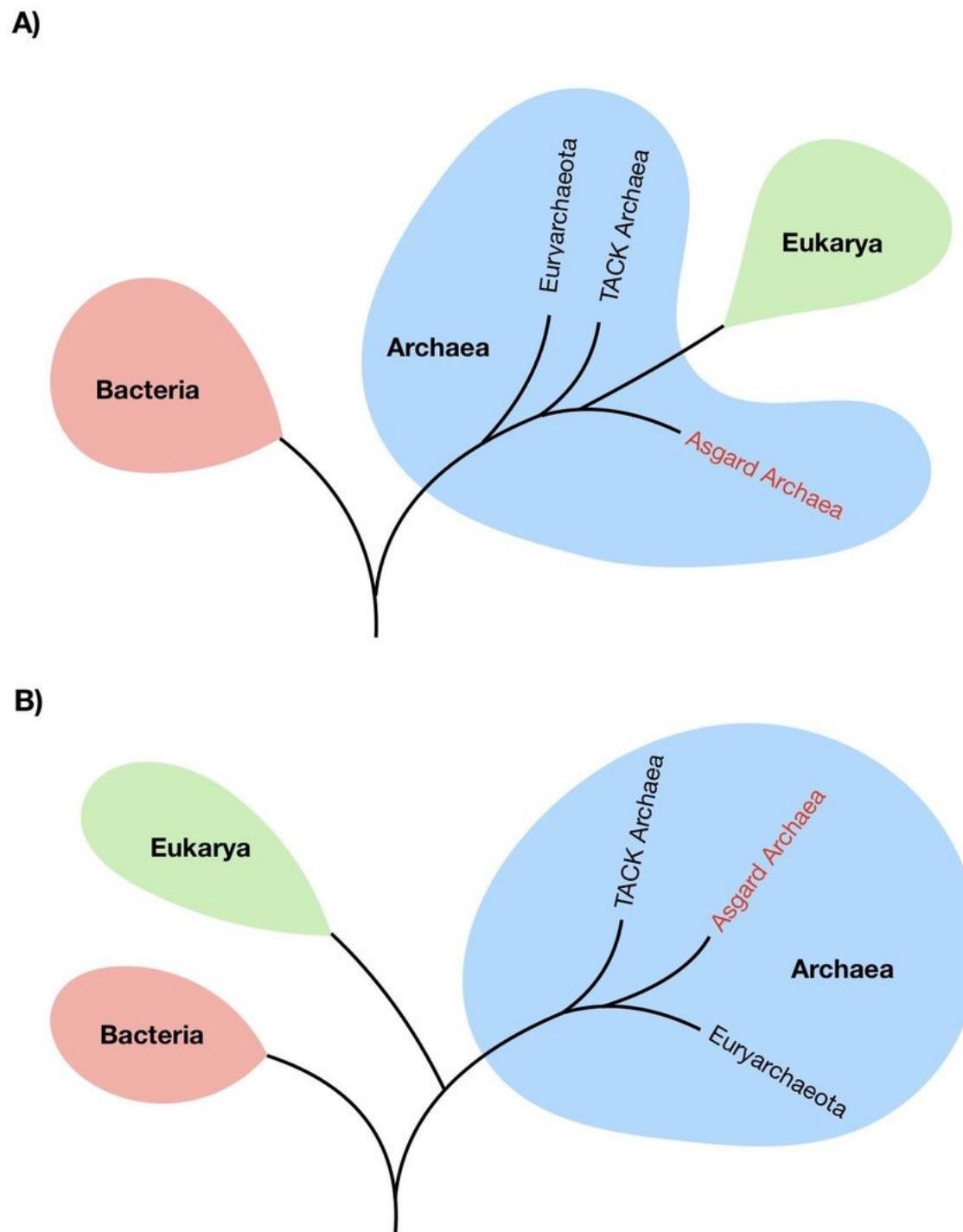


The Eocyte hypothesis

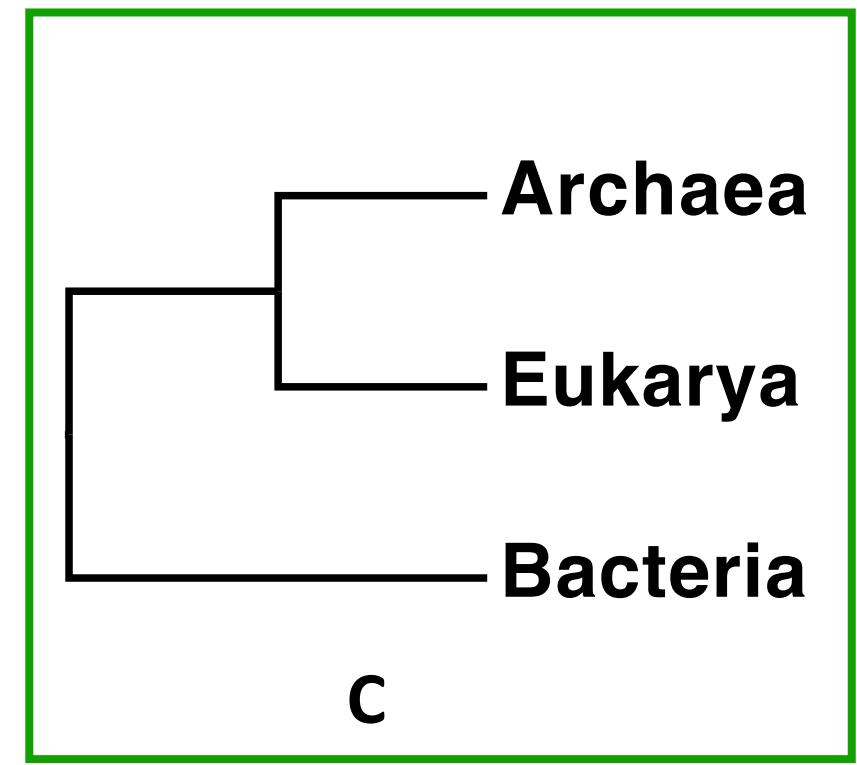
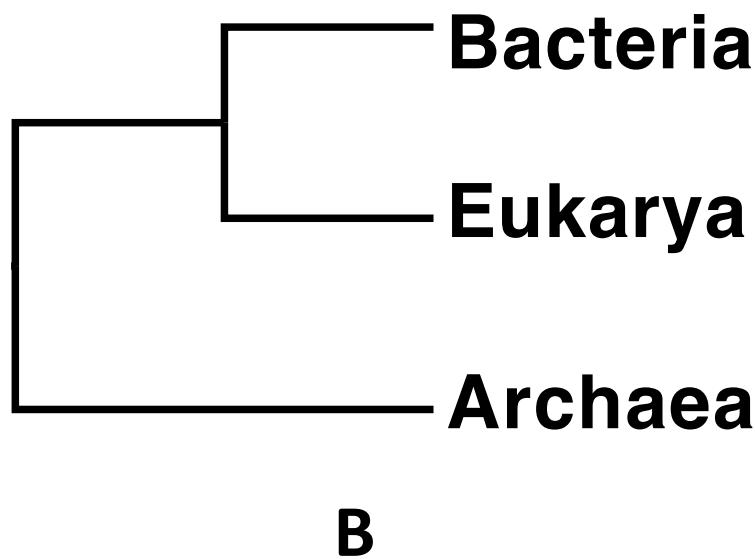
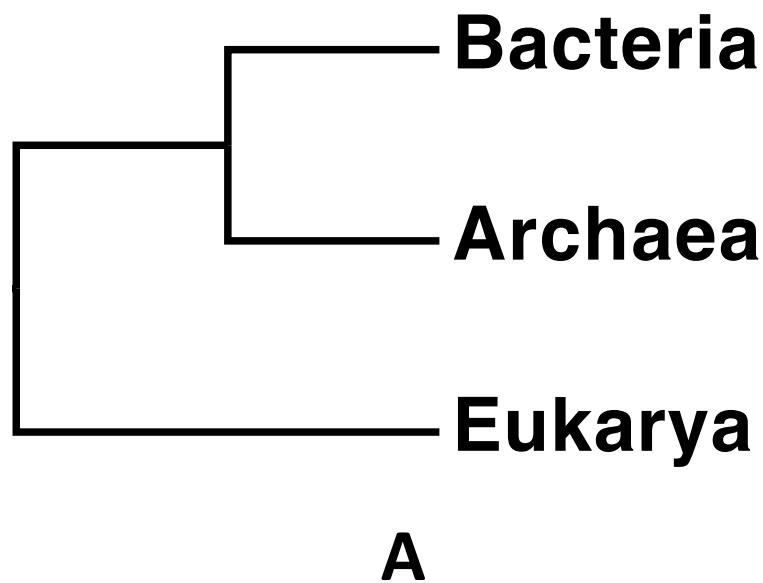
# Origin of Eukarya

- Which lineage is the sister lineage to Eukarya?
- Discovery of Asgard Archaea is complicating the picture

Figure 5. (A) The tree of life resulting from a range of phylogenetic analyses of **conserved markers**, **ribosomal RNA genes** and **ribosomal proteins**, placing Asgard archaea as the closest archaeal relatives of eukaryotes. (B) The tree produced from phylogenetic analyses of **Asgard RNA polymerase genes** supports the three-domain topology of the tree of life, with Asgard archaea as a sister group to Euryarchaeota.



Best hypothesis for the relationships among the three domains of life is:



# Molecular evolution: arms race between caterpillars and their host plants

- Plants have secondary metabolites
- One of the functions: protection from herbivory
- Insects can evolve the ability to metabolize plant secondary metabolites
- Then plants can evolve new chemicals, and we have an arms race
- Example: Edger et al. 2015 PNAS study



## The butterfly plant arms-race escalated by gene and genome duplications

Patrick P. Edger<sup>a,b,c,1</sup>, Hanna M. Heidel-Fischer<sup>d,1</sup>, Michaël Bekaert<sup>e</sup>, Jadranka Rota<sup>f</sup>, Gernot Glöckner<sup>g,h</sup>, Adrian E. Platts<sup>i</sup>, David G. Heckel<sup>d</sup>, Joshua P. Der<sup>j,k</sup>, Eric K. Wafula<sup>j</sup>, Michelle Tang<sup>a</sup>, Johannes A. Hofberger<sup>l</sup>, Ann Smithson<sup>m,n</sup>, Jocelyn C. Hall<sup>o</sup>, Matthieu Blanchette<sup>i</sup>, Thomas E. Bureau<sup>p</sup>, Stephen I. Wright<sup>q</sup>, Claude W. dePamphilis<sup>j</sup>, M. Eric Schranz<sup>j</sup>, Michael S. Barker<sup>b</sup>, Gavin C. Conant<sup>r,s</sup>, Niklas Wahlberg<sup>f</sup>, Heiko Vogel<sup>d</sup>, J. Chris Pires<sup>a,s,2</sup>, and Christopher W. Wheat<sup>t,2</sup>

<sup>a</sup>Division of Biological Sciences, University of Missouri, Columbia, MO 65211; <sup>b</sup>Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721; <sup>c</sup>Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720; <sup>d</sup>Department of Entomology, Max Planck Institute for Chemical Ecology, 07745 Jena, Germany; <sup>e</sup>Institute of Aquaculture, University of Stirling, Stirling FK9 4LA, Scotland, United Kingdom;

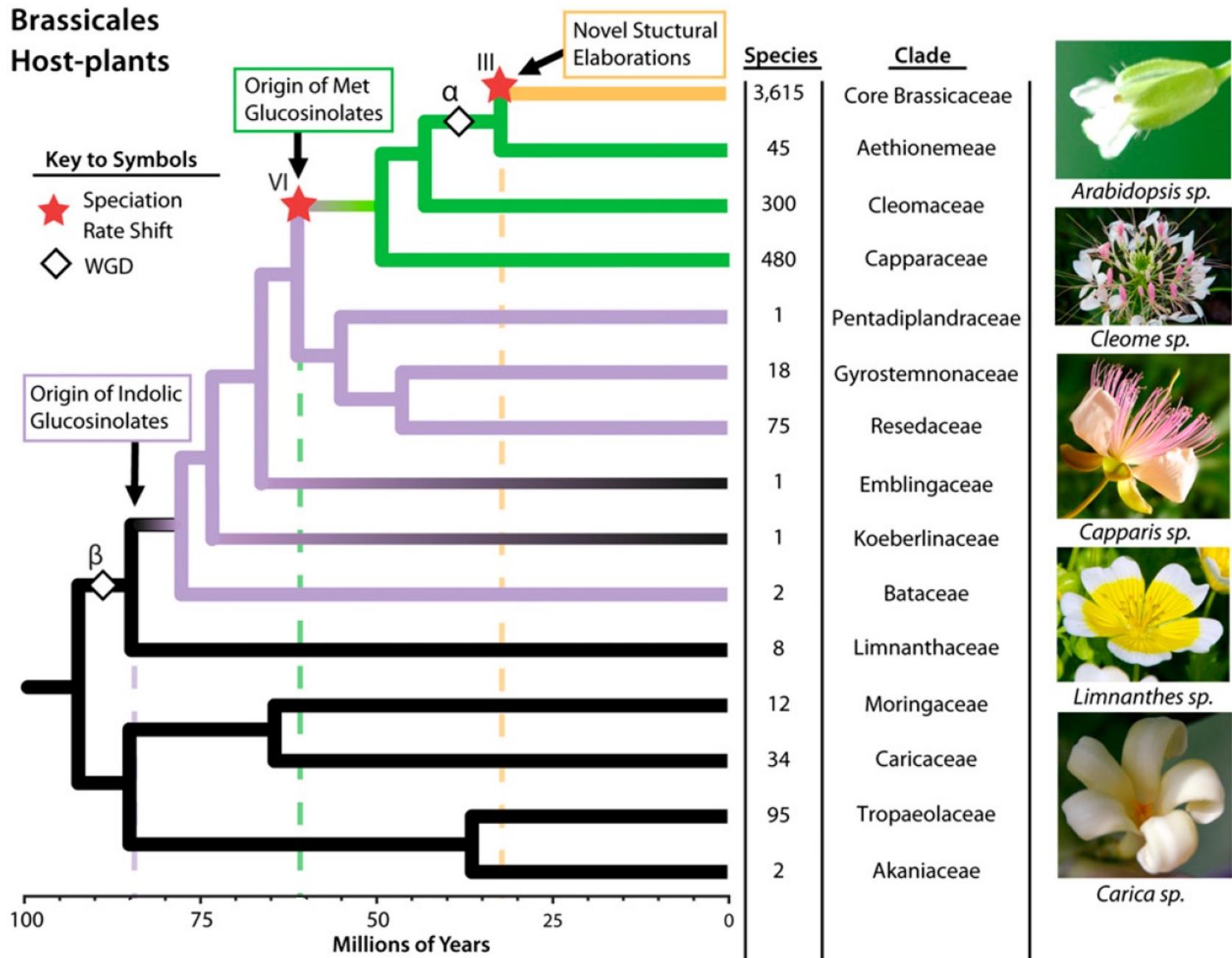
# Edger et al. 2015. The butterfly plant arms-race escalated by gene and genome duplications

- Butterflies: family Pieridae
- Plants: order Brassicales
- A number of gradual changes
- BUT also gene and genome duplications
- The story: “Nearly 90 million years ago, the ancestors of *Brassica* (mustards, cabbage) and related plants developed a chemical defense called glucosinolates.”
  - Glucosinolates – toxic to most insects, for humans source of the sharp taste in wasabi, horseradish, and mustard

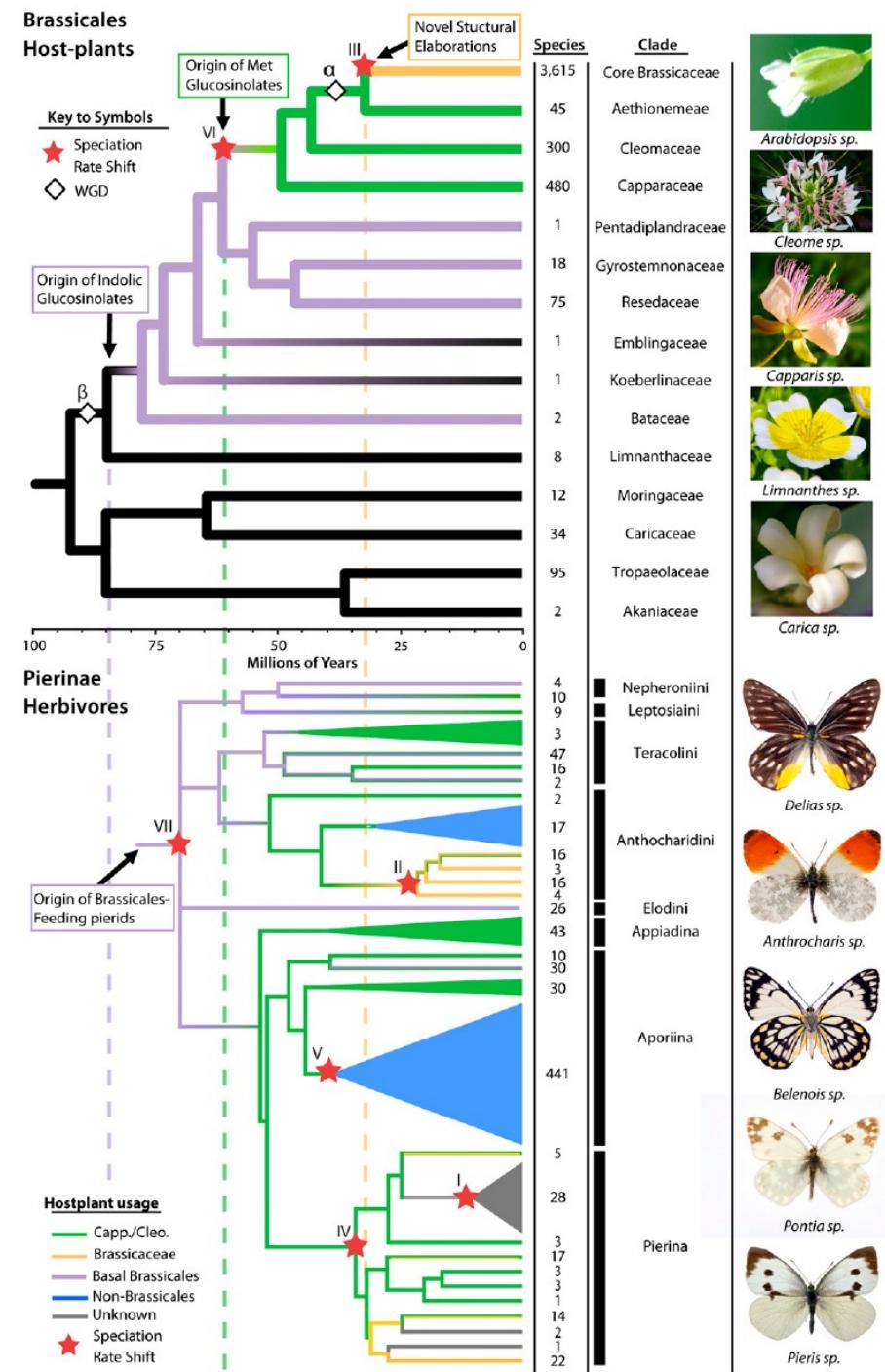


Cabbage Butterfly *Pieris brassicae* (Linnaeus, 1758)  
By Didier Descouens - Own work, CC BY-SA 4.0

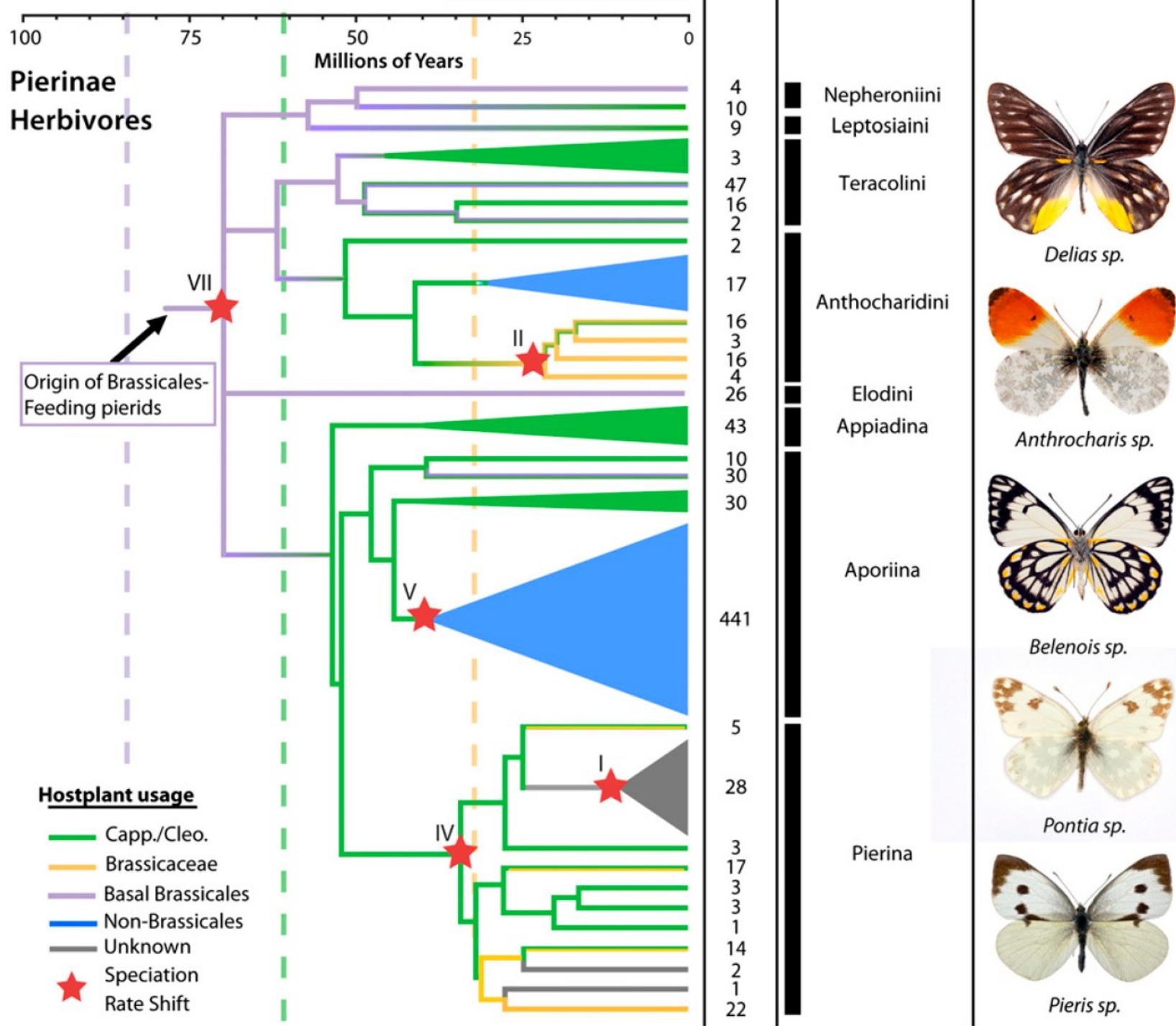
**WGD**  
= whole genome duplications



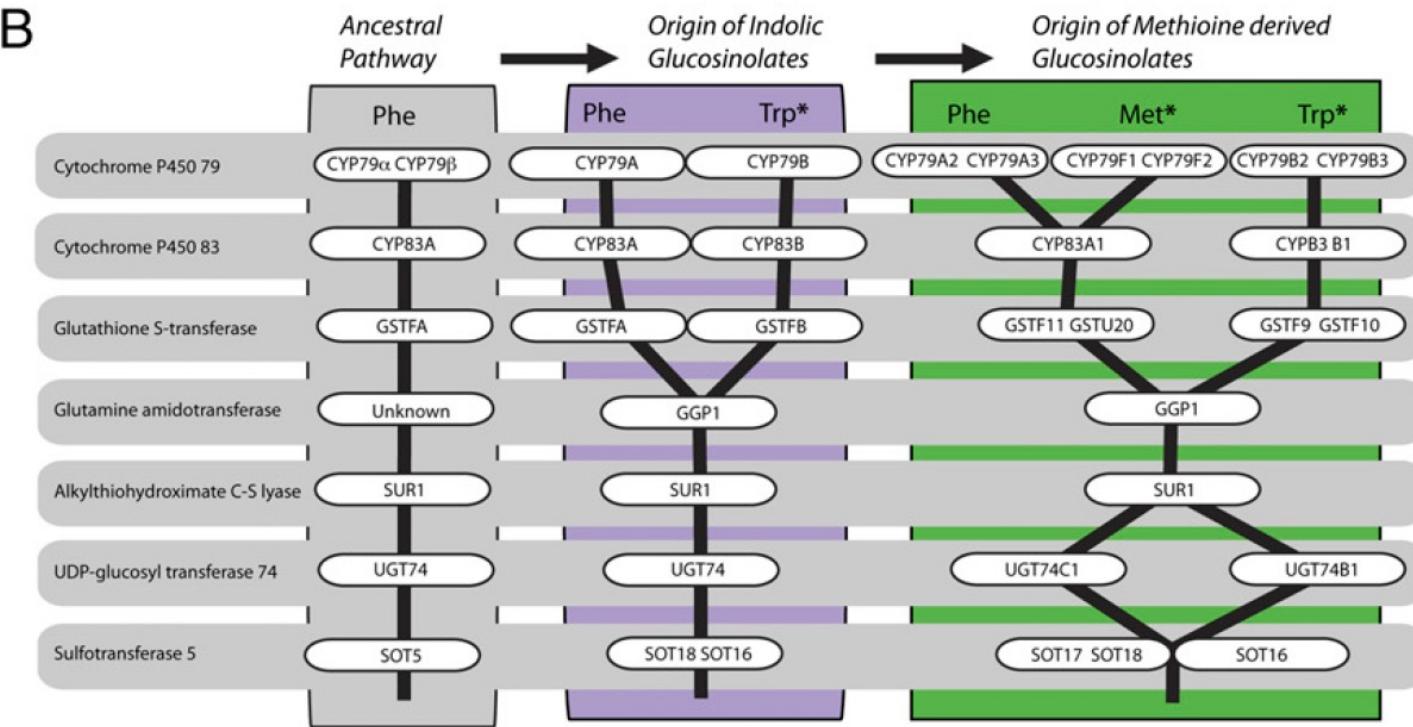
- Timing of divergence for plants and butterflies carried out on independent data
- Butterflies colonized the host plants after they had evolved their chemical defences



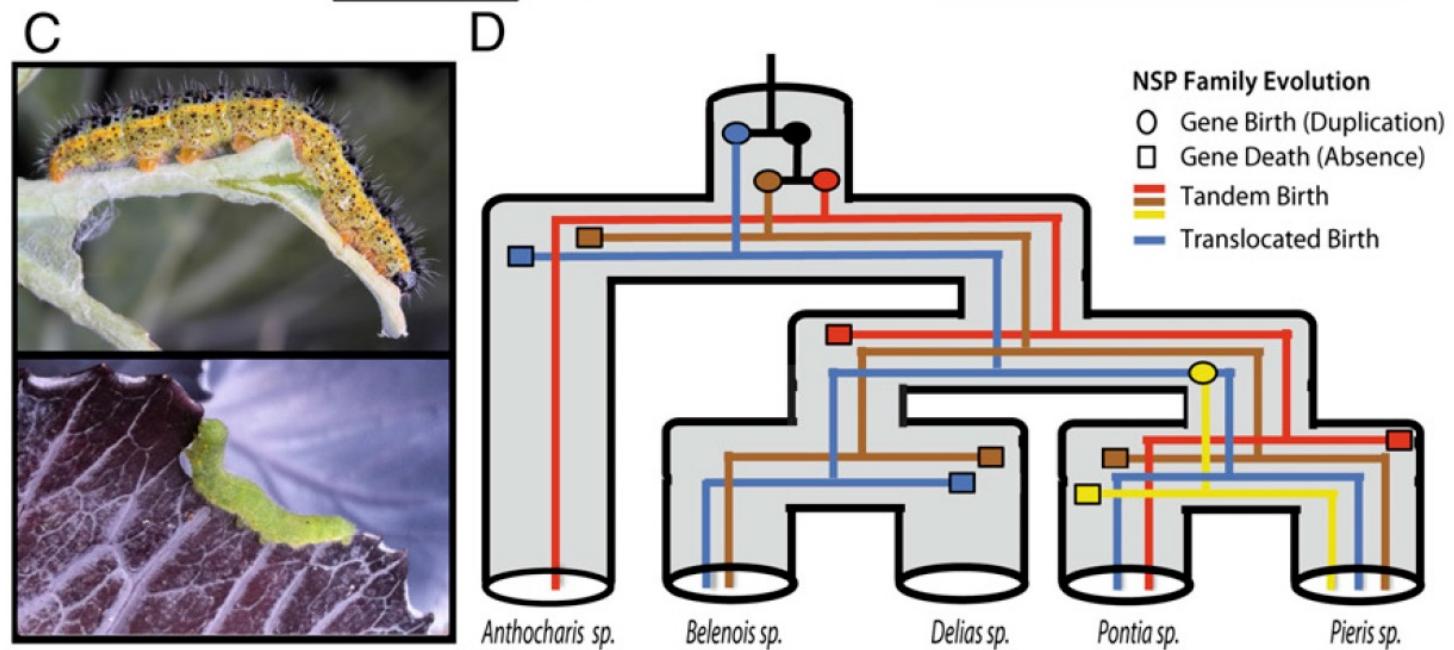
- It appears that colonizing chemically protected host plants allowed the butterflies to speciate faster



**Figure 2. (B)** An illustration of the evolution of core glucosinolate pathways across Brassicales; with substrates tryptophan (Trp), phenylalanine (Phe), and methionine (Met) shown at the top, enzymes depicted as white ovals, and each pathway as black vertical lines.



**Figure 2. (D)** Evolution of the NSP gene family is shown across select Pierinae genera, indicating the birth and death dynamics of four paralogous clades.

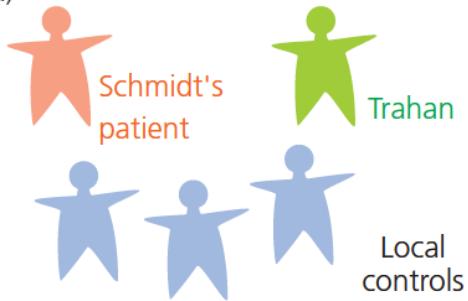


\***Nitrile Specifier Protein (NSP):** primary detoxification mechanism used by the butterflies to break down this chemical defense system

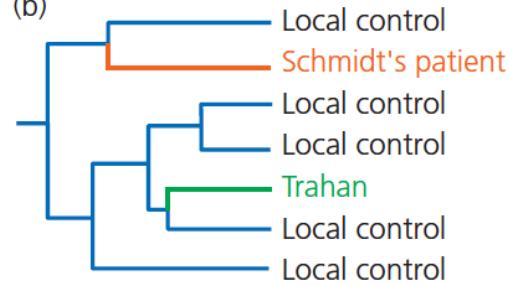
# Criminal investigations: evolutionary history of a deliberate HIV infection

- Source: Herron & Freeman “Evolutionary Analysis” 5<sup>th</sup> edition
- Players in this true story:
  - Janet Trahan, nurse – the accuser
  - Richard Schmidt, physician – the accused
  - David Hillis, evolutionary biologist – providing crucial evidence
- Legal case:
  - Trahan became sick with HIV and accused Schmidt of deliberately injecting her with blood from a HIV-positive patient as revenge for breaking off their affair
- Analysis:
  - Evolutionary history of viral samples from various people to establish from where Trahan contracted her virus

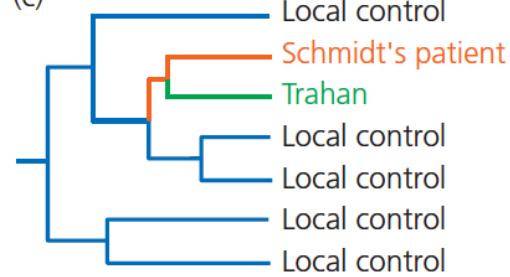
(a)



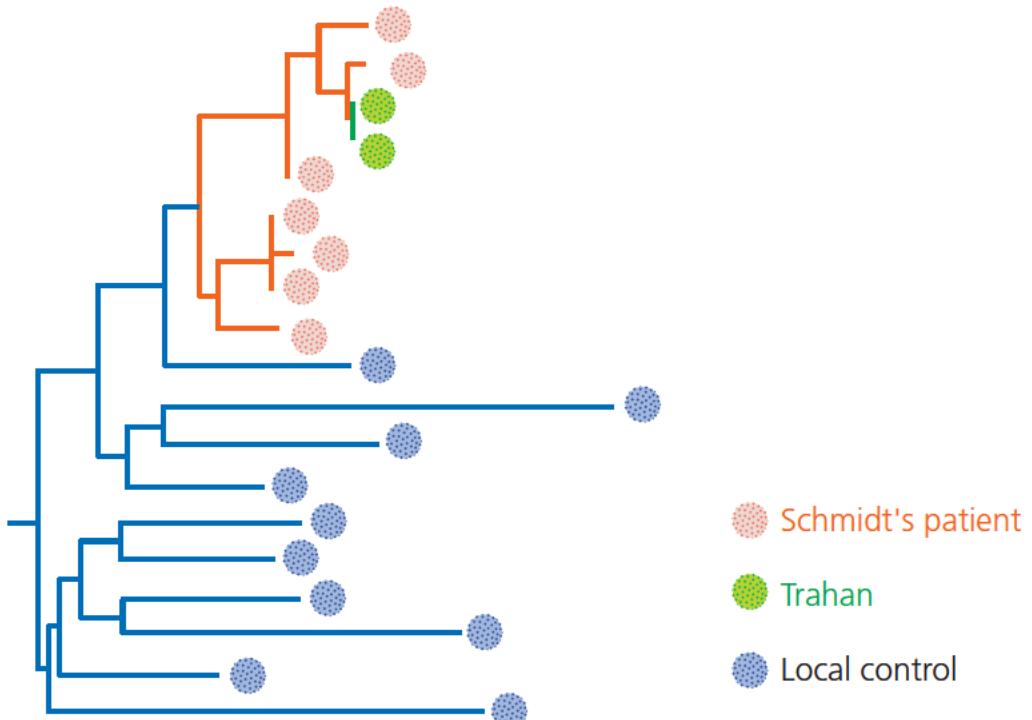
(b)



(c)



**Figure 1.23 Reconstructing evolutionary history tests the prosecution's theory of a crime** (a) The individuals from whom HIV samples were collected. (b) The viral evolutionary tree predicted if the suspect is innocent. (c) The evolutionary tree predicted if the suspect is guilty.



**Figure 1.24 Incriminating evidence** This reconstructed evolutionary history is consistent with the prosecution's charges. Redrawn from Metzker et al. (2002).

# **Importance of Understanding Evolutionary History:**

examples we talked about (there could be many, many more)

- **Viral evolution and its impact on human health**
- **Systematics – understanding the relationships among different groups of organisms and how they acquired their current form and function**
  - E.g. snakes and lizards, birds and dinosaurs, the main domains of life
- **Molecular evolution**
  - Genome and gene duplications as weapons in an arms race between plants and caterpillars
- **Even criminal investigations!**