



Tree-thinking and introduction to phylogenetics

Nothing in biology makes sense except in the light of evolution.

— Theodosius Dobzhansky, 1973

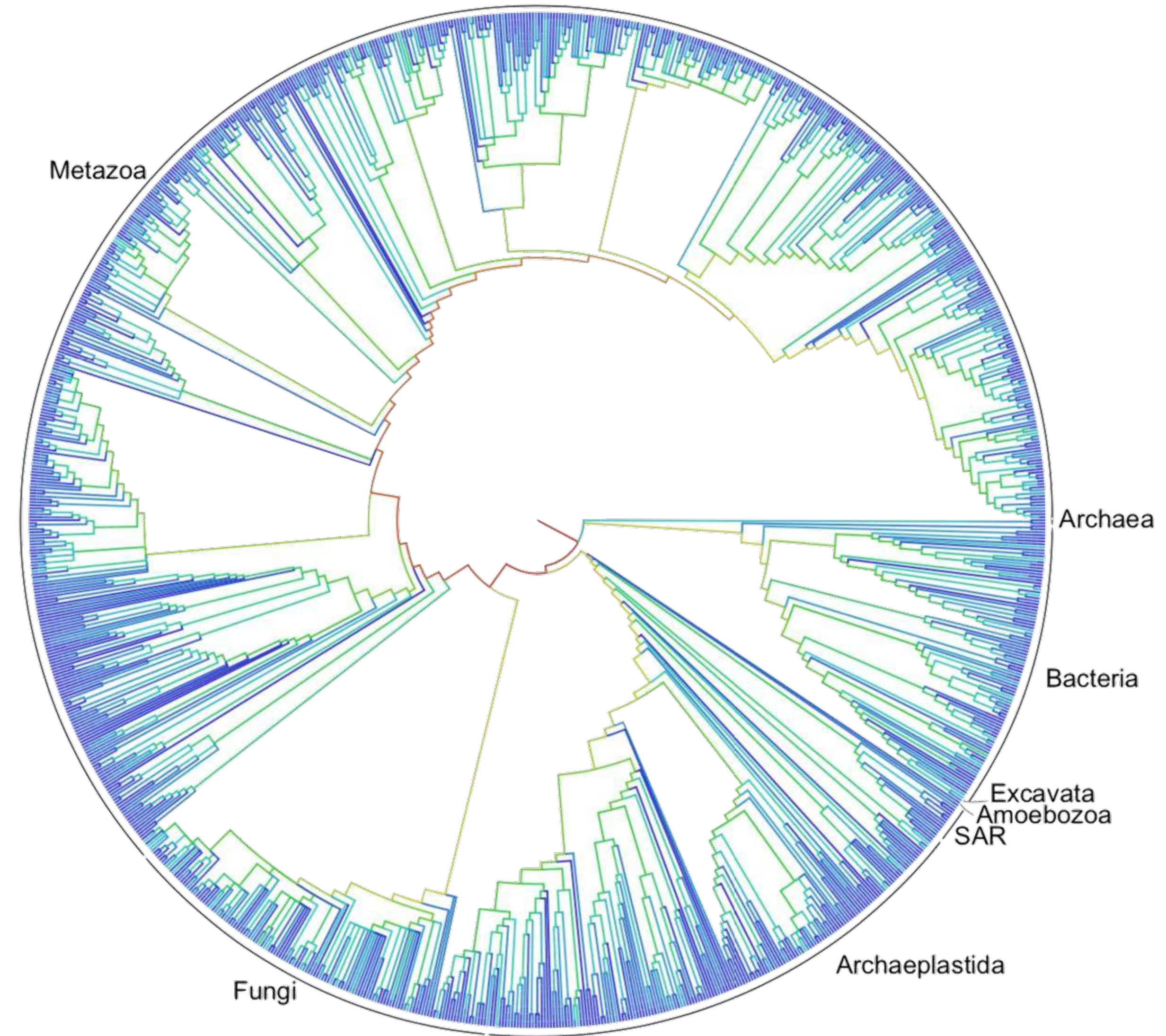
Nothing in evolution makes sense except when seen in the light of phylogeny.

— Jay Savage, 1997

The Tree of Life

*Three billion years the Tree has grown
From replicators' first seed sown
To branches rich with progeny:
The wonder of phylogeny.*

excerpt from the poem "[The Tree of Life](#)" by David Maddison



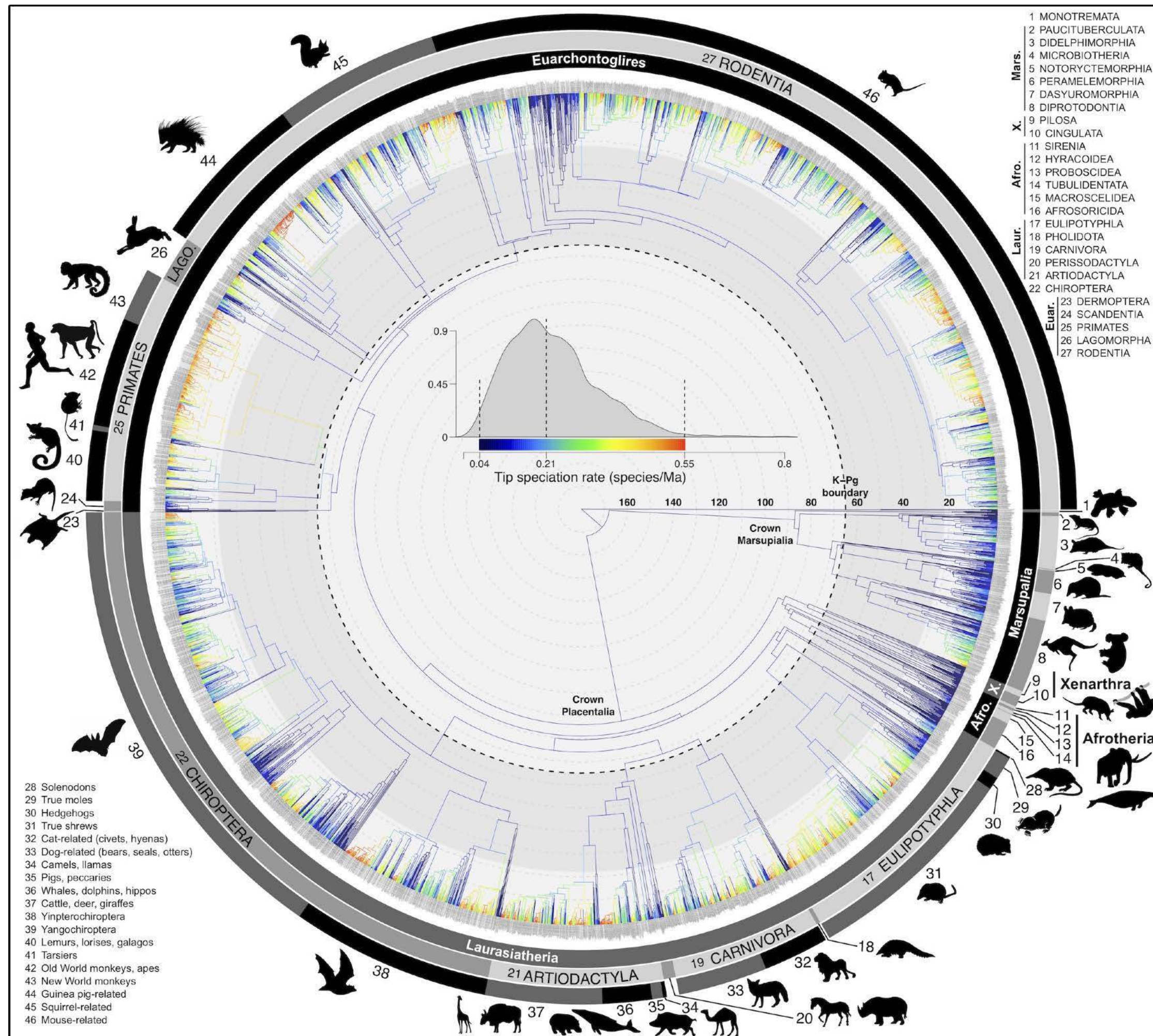
Systematics and Phylogenies

The study of biodiversity

How are species delimited?

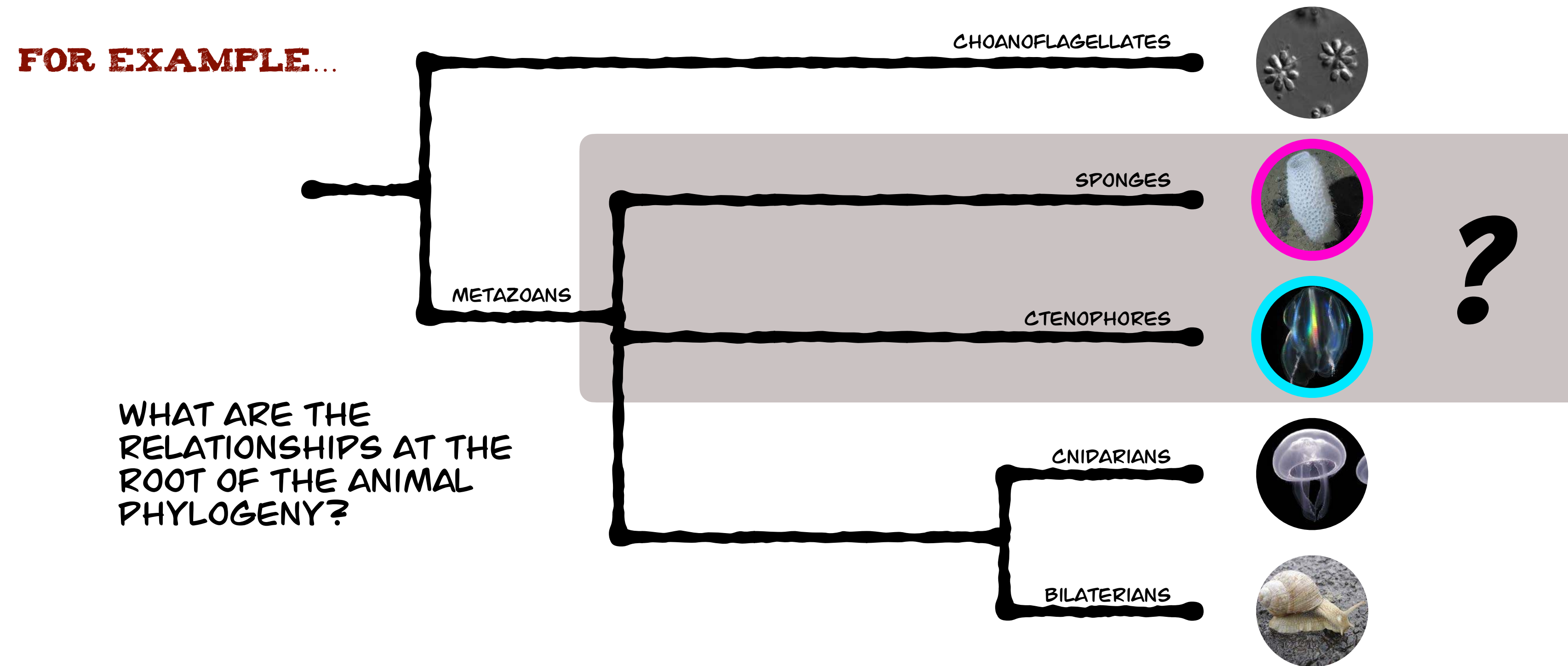
How are species organized and distributed?

What evolutionary and ecological processes produced the patterns of biodiversity observed on Earth?



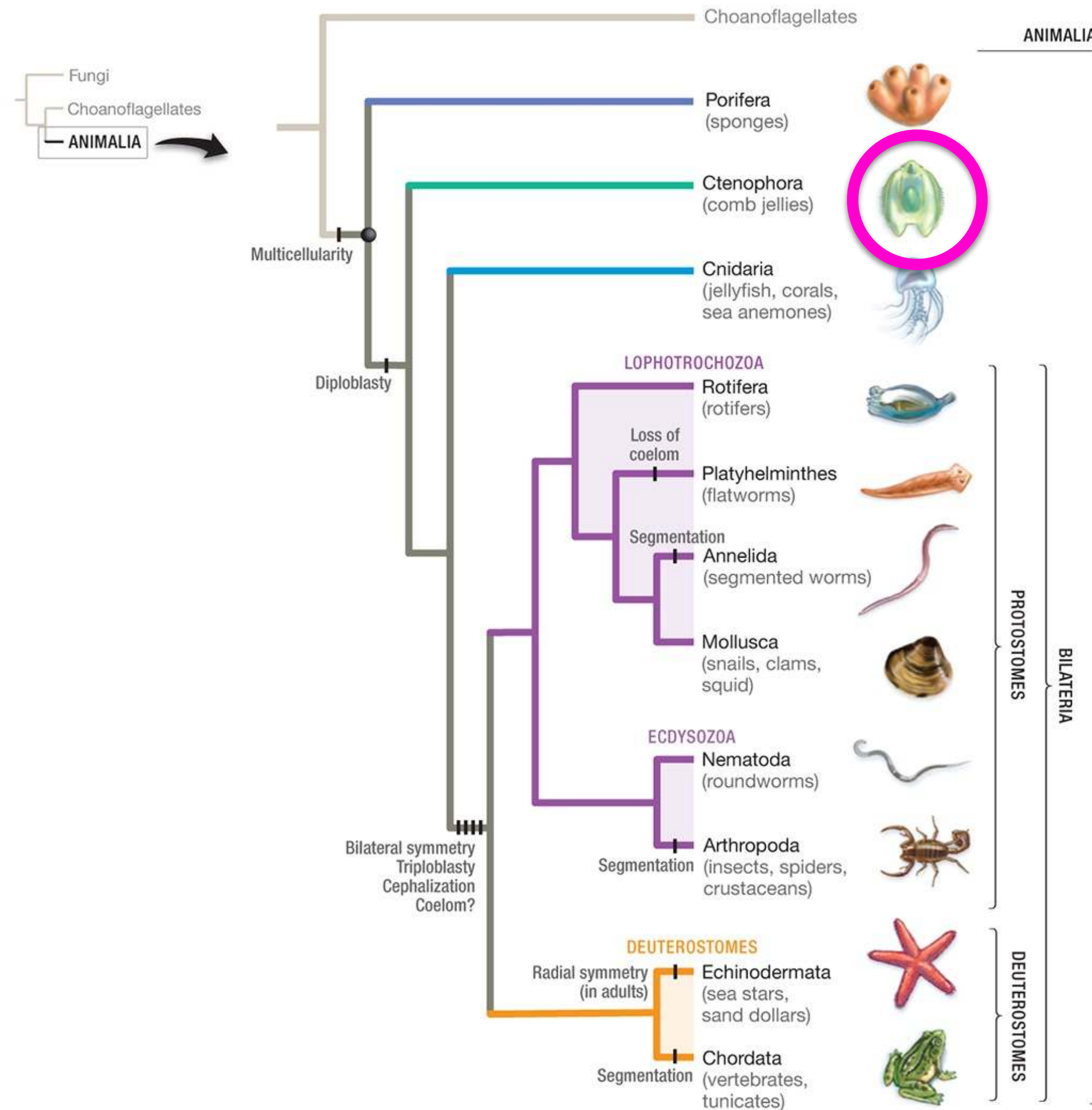
“Resolving” the Tree of Life

We can never know the *true* phylogeny of life on Earth...



...but we can work to consider all of the data and use rigorous inference methods to produce reliable estimates.

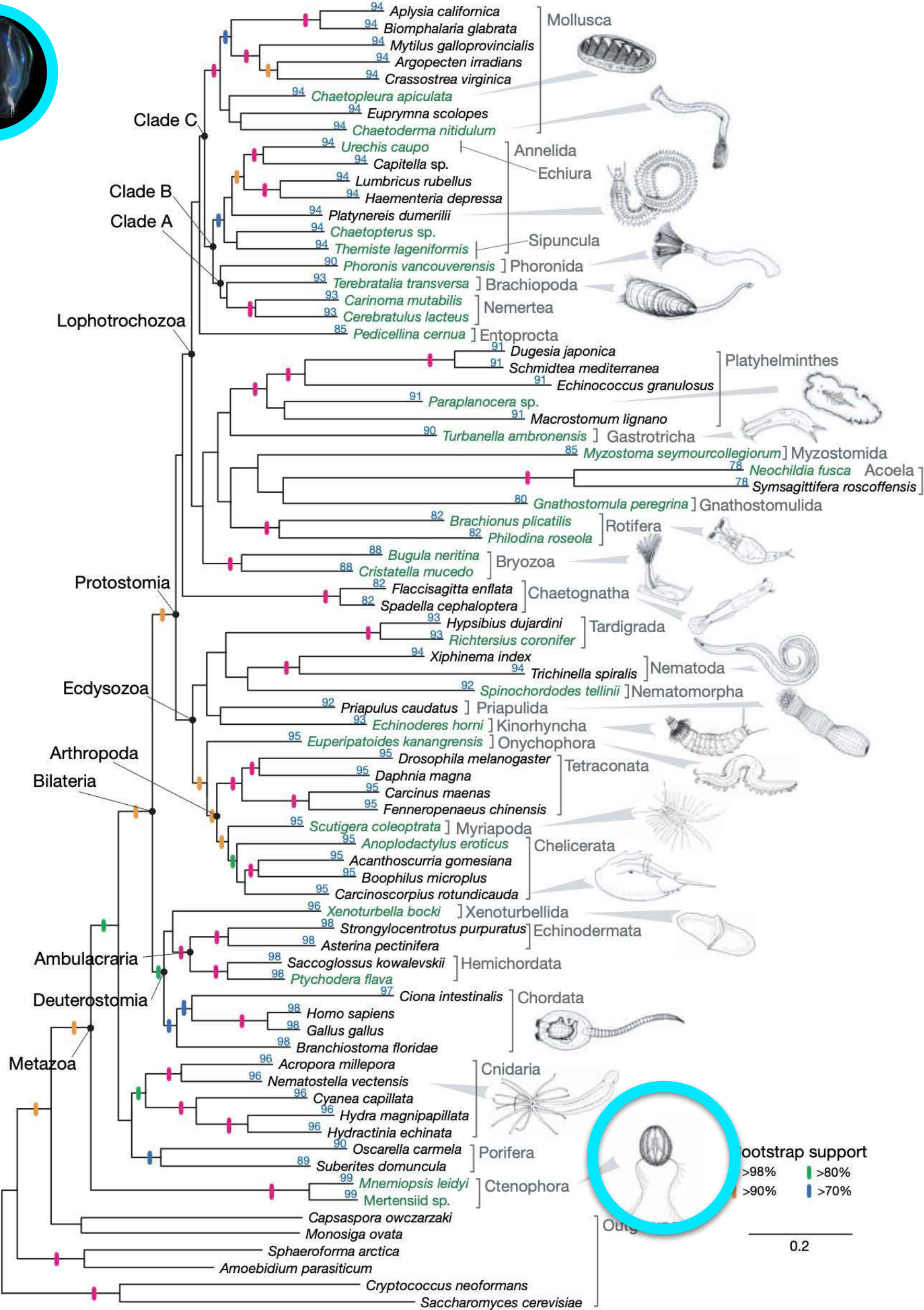
The Animal Phylogeny



Until recently, there was relatively little debate about the relationships at the root of the animal tree

Morphological evidence indicated that sponges were sister to all other animal lineages

Ctenophores Spark Debate



Letter | Published: 05 March 2008

Broad phylogenomic sampling improves resolution of the animal tree of life

[Casey W. Dunn](#) , [Andreas Hejnol](#), [David Q. Matus](#), [Kevin Pang](#), [William E. Browne](#), [Stephen A. Smith](#), [Elaine Seaver](#), [Greg W. Rouse](#), [Matthias Obst](#), [Gregory D. Edgecombe](#), [Martin V. Sørensen](#), [Steven H. D. Haddock](#), [Andreas Schmidt-Rhaesa](#), [Akiko Okusu](#), [Reinhardt Møbjerg Kristensen](#), [Ward C. Wheeler](#), [Mark Q. Martindale](#) & [Gonzalo Giribet](#)

[Nature](#) **452**, 745–749 (2008) | [Cite this article](#)



Who's my sister?

PNAS

Error, signal, and the placement of Ctenophora sister to all other animals

Nathan V. Whelan^{a,1}, Kevin M. Kocot^b, Leonid L. Moroz^c, and Kenneth M. Halanych^a

April 20, 2015 | 112 (18) 5773-5778 | <https://doi.org/10.1073/pnas.1503453112>



Research article | [Open access](#) | Published: 23 November 2015

Extracting phylogenetic signal and accounting for bias in whole-genome data sets supports the Ctenophora as sister to remaining Metazoa

[Marek L. Borowiec](#) , [Ernest K. Lee](#), [Joanna C. Chiu](#) & [David C. Plachetzki](#) 

[BMC Genomics](#) **16**, Article number: 987 (2015) | [Cite this article](#)



ARTICLE · Volume 27, Issue 7, P958-967, April 03, 2017 ·

Current Biology
Article

A Large and Consistent Phylogenomic Dataset Supports Sponges as the Sister Group to All Other Animals

[Paul Simion](#) ^{1,17,18} · [Hervé Philippe](#) ^{2,3,17}   · [Denis Baurain](#) ⁴ · [Muriel Jager](#) ¹ · [Daniel J. Richter](#) ^{5,6,7} · [Arnaud Di Franco](#) ² · [Béatrice Roure](#) ^{2,3} · [Nori Satoh](#) ⁸ · [Éric Quéinnec](#) ¹ · [Alexander Ereskovsky](#) ^{9,10} · [Pascal Lapébie](#) ¹¹ · [Erwan Corre](#) ^{12,13} · [Frédéric Delsuc](#) ¹⁴ · [Nicole King](#) ⁵ · [Gert Wörheide](#) ^{15,16} · [Michaël Manuel](#) ^{1,19} 



Contentious relationships in phylogenomic studies can be driven by a handful of genes

[Xing-Xing Shen](#), [Chris Todd Hittinger](#) & [Antonis Rokas](#) 

[Nature Ecology & Evolution](#) **1**, Article number: 0126 (2017) | [Cite this article](#)



Letter | Published: 05 March 2008

Broad phylogenomic sampling improves resolution of the animal tree of life

[Casey W. Dunn](#) , [Andreas Hejnol](#), [David Q. Matus](#), [Kevin Pang](#), [William E. Browne](#), [Stephen A. Smith](#), [Elaine Seaver](#), [Greg W. Rouse](#), [Matthias Obst](#), [Gregory D. Edgecombe](#), [Martin V. Sørensen](#), [Steven H. D. Haddock](#), [Andreas Schmidt-Rhaesa](#), [Akiko Okusu](#), [Reinhardt Møbjerg Kristensen](#), [Ward C. Wheeler](#), [Mark Q. Martindale](#) & [Gonzalo Giribet](#)

[Nature](#) **452**, 745–749 (2008) | [Cite this article](#)



PNAS

Genomic data do not support comb jellies as the sister group to all other animals

[Davide Pisani](#) ^{a,b,1}, [Walker Pett](#) ^c, [Martin Dohrmann](#) ^d, [Roberto Feuda](#) ^e, [Omar Rota-Stabelli](#) ^f, [Hervé Philippe](#) ^{g,h}, [Nicolas Lartillot](#) ^c, and [Gert Wörheide](#) ^{d,i,1}

November 30, 2015 | 112 (50) 15402-15407 | <https://doi.org/10.1073/pnas.1518127112>



The Guardian

Evolution row ends as scientists declare sponges to be sister of all other animals

Nicola Davis

Thu 30 Nov 2017 12:00 EST



REPORT · Volume 27, Issue 24, P3864-3870.E4, December 18, 2017 ·

Current Biology
Report

Improved Modeling of Compositional Heterogeneity Supports Sponges as Sister to All Other Animals

[Roberto Feuda](#) ^{1,11} · [Martin Dohrmann](#) ^{2,11} · [Walker Pett](#) ³ · [Hervé Philippe](#) ^{4,5} · [Omar Rota-Stabelli](#) ⁶ · [Nicolas Lartillot](#) ⁷ · [Gert Wörheide](#) ^{2,8,9}   · [Davide Pisani](#) ^{10,12} 



Evidence for sponges as sister to all other animals from partitioned phylogenomics with mixture models and recoding

[Anthony K. Redmond](#) & [Aoife McLysaght](#) 

[Nature Communications](#) **12**, Article number: 1783 (2021) | [Cite this article](#)



New Evidence

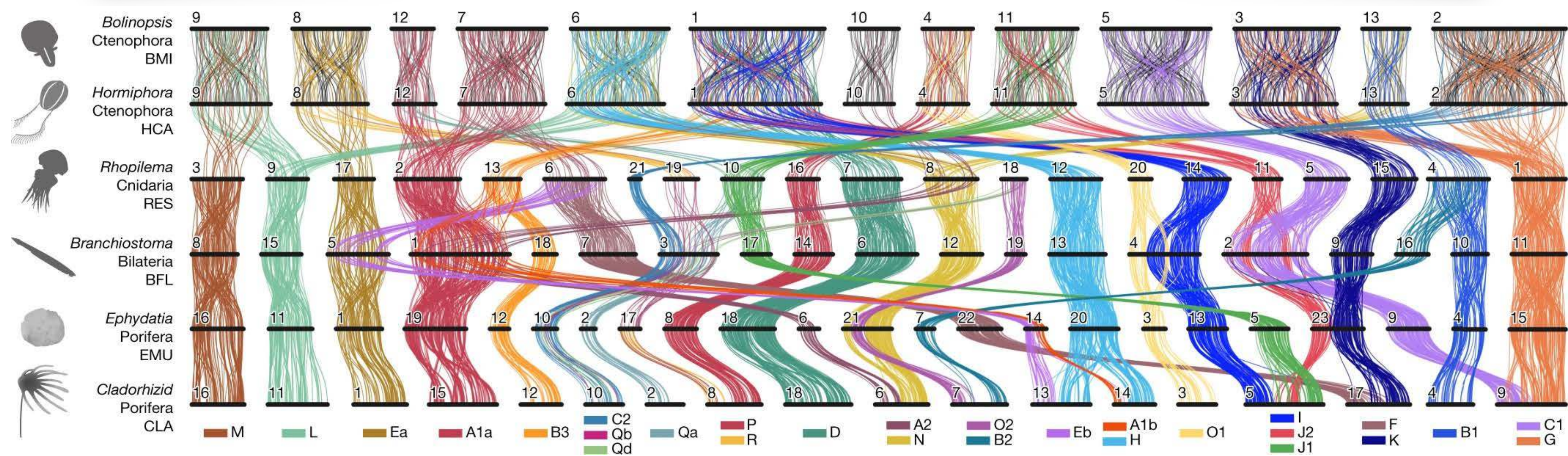


Article | [Open access](#) | Published: 17 May 2023

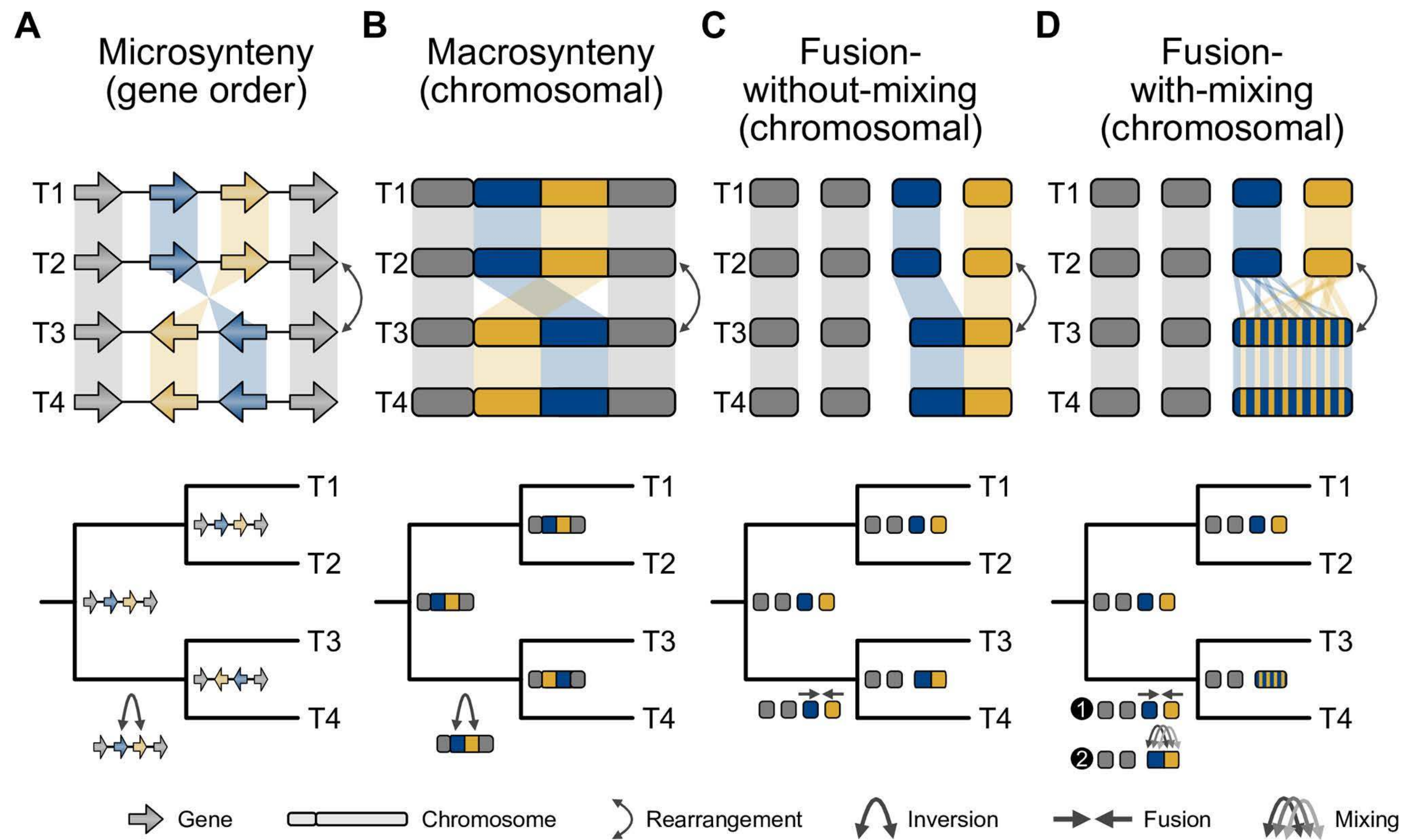
Ancient gene linkages support ctenophores as sister to other animals

[Darrin T. Schultz](#) , [Steven H. D. Haddock](#), [Jessen V. Bredeson](#), [Richard E. Green](#), [Oleg Simakov](#)  & [Daniel S. Rokhsar](#) 

[Nature](#) **618**, 110–117 (2023)



Large-Scale Gene Order Changes Are Rare



Conserved Synteny

Article | [Open access](#) | Published: 17 May 2023

Ancient gene linkages support ctenophores as sister to other animals

[Darrin T. Schultz](#) , [Steven H. D. Haddock](#), [Jessen V. Bredeson](#), [Richard E. Green](#), [Oleg Simakov](#)  & [Daniel S. Rokhsar](#) 

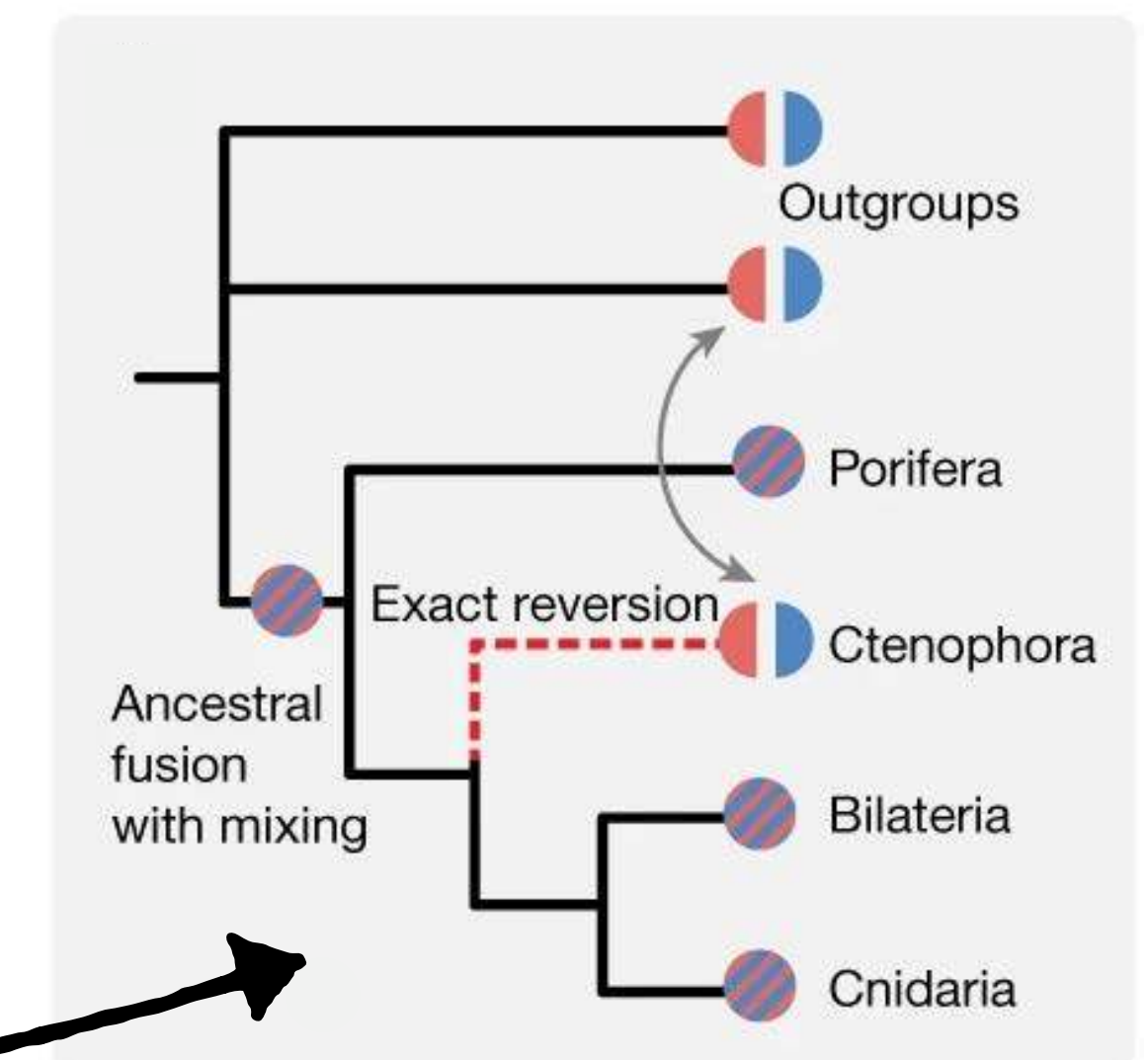
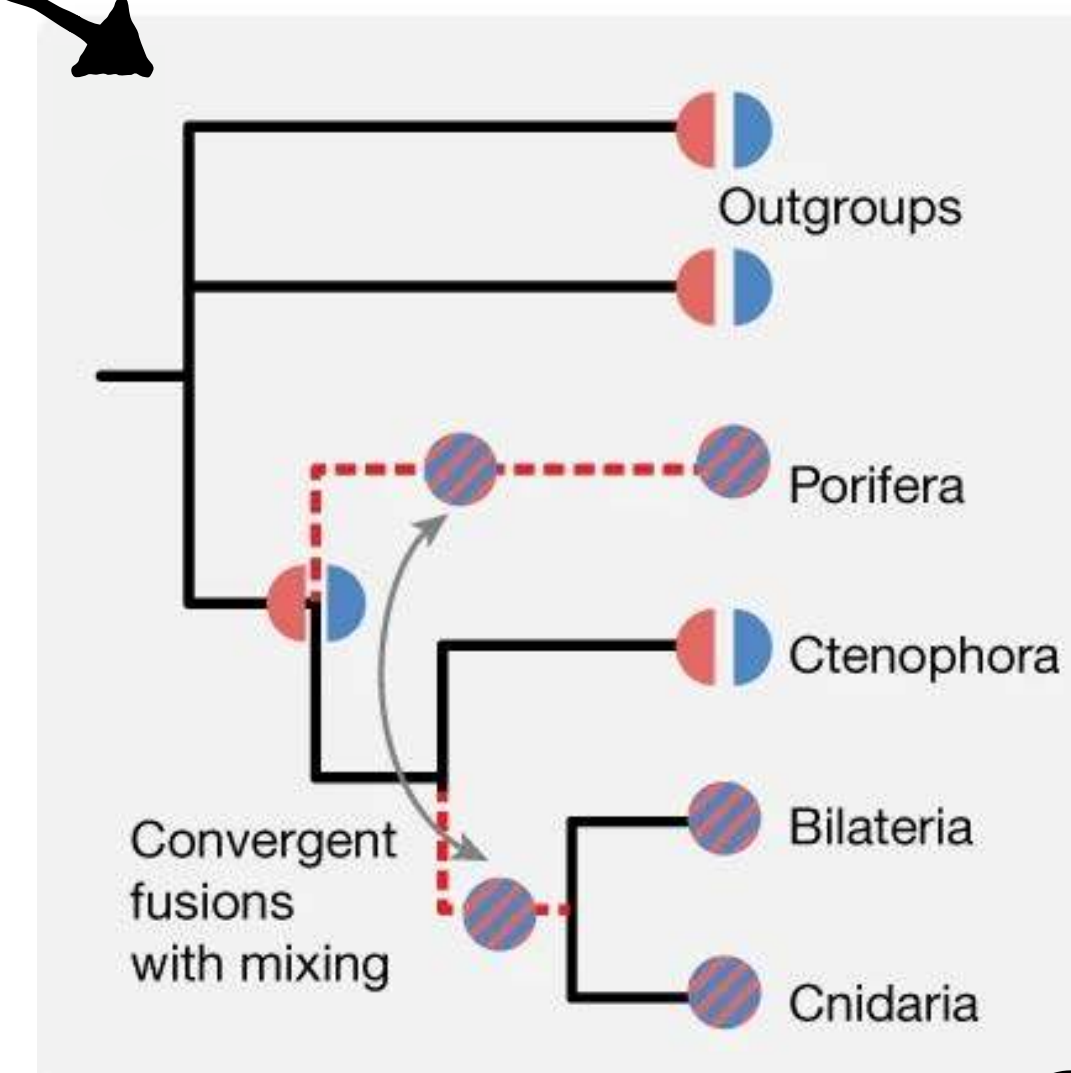
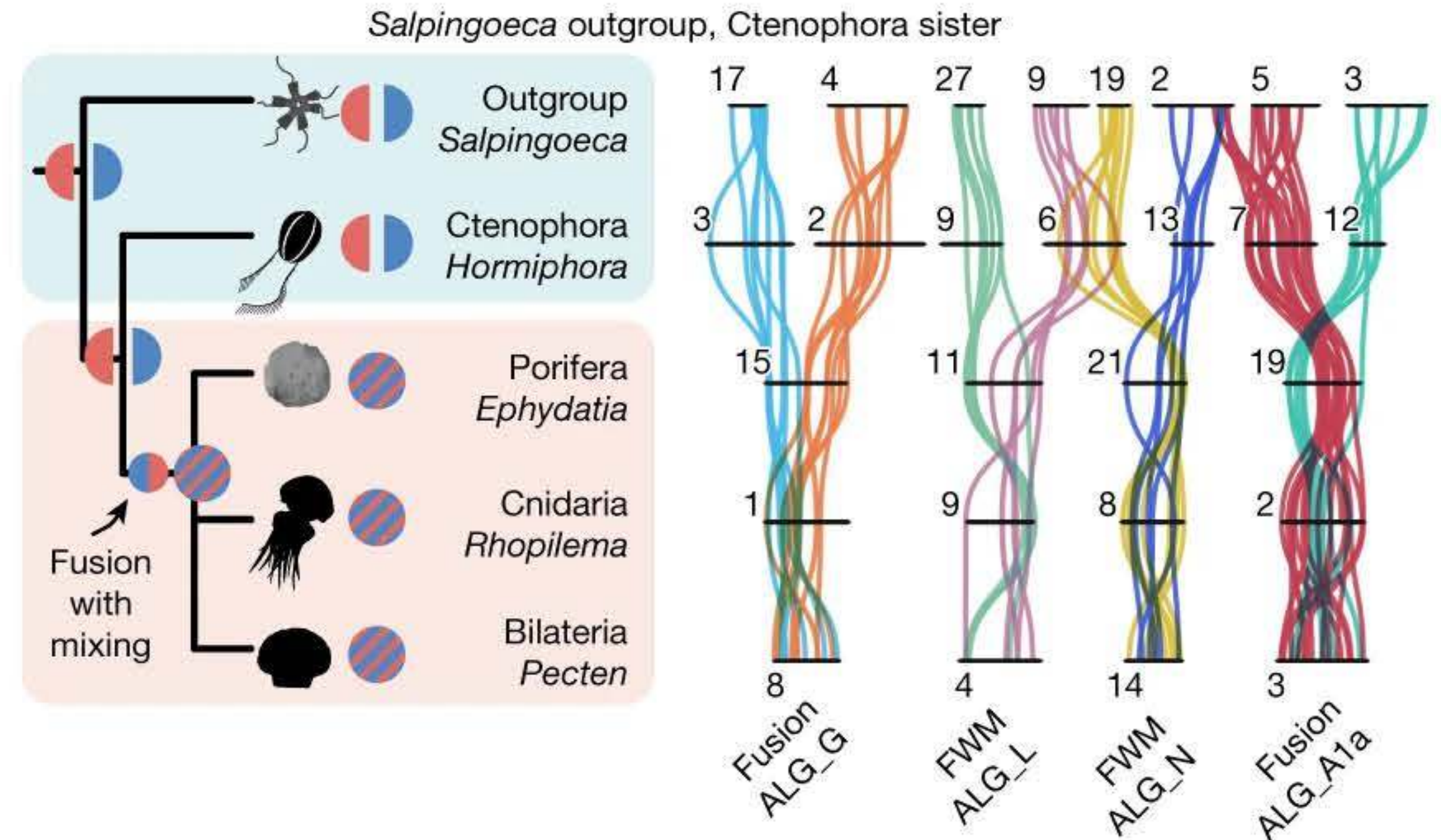
[Nature](#) **618**, 110–117 (2023)

Unlikely changes are needed for these patterns to exist if sponges are sister to all other animals

Convergent fusions followed by mixing

or

Exact unmixing followed by fissions



Phylogenies Give Evolutionary Context

Understanding the complex history of animals requires more research on ctenophore & sponge biology since these groups aren't well studied & often misrepresented

The evolution of the neuromuscular system is more complicated than we originally assumed

Syncytial nerve net in a ctenophore adds insights on the evolution of nervous systems

PAWEŁ BURKHARDT , JEFFREY COLGREN , ASTRID MEDHUS , LEONID DIGEL , BENJAMIN NAUMANN , JOAN J. SOTO-ANGEL , EVA-LENA NORDMANN 
, MARIA Y. SACHKOVA , AND MAIKE KITTELMANN  [Authors Info & Affiliations](#)

SCIENCE • 20 Apr 2023 • Vol 380, Issue 6642 • pp. 293-297 • DOI: 10.1126/science.ade5645

And new data and/or methods may lead to new insights...

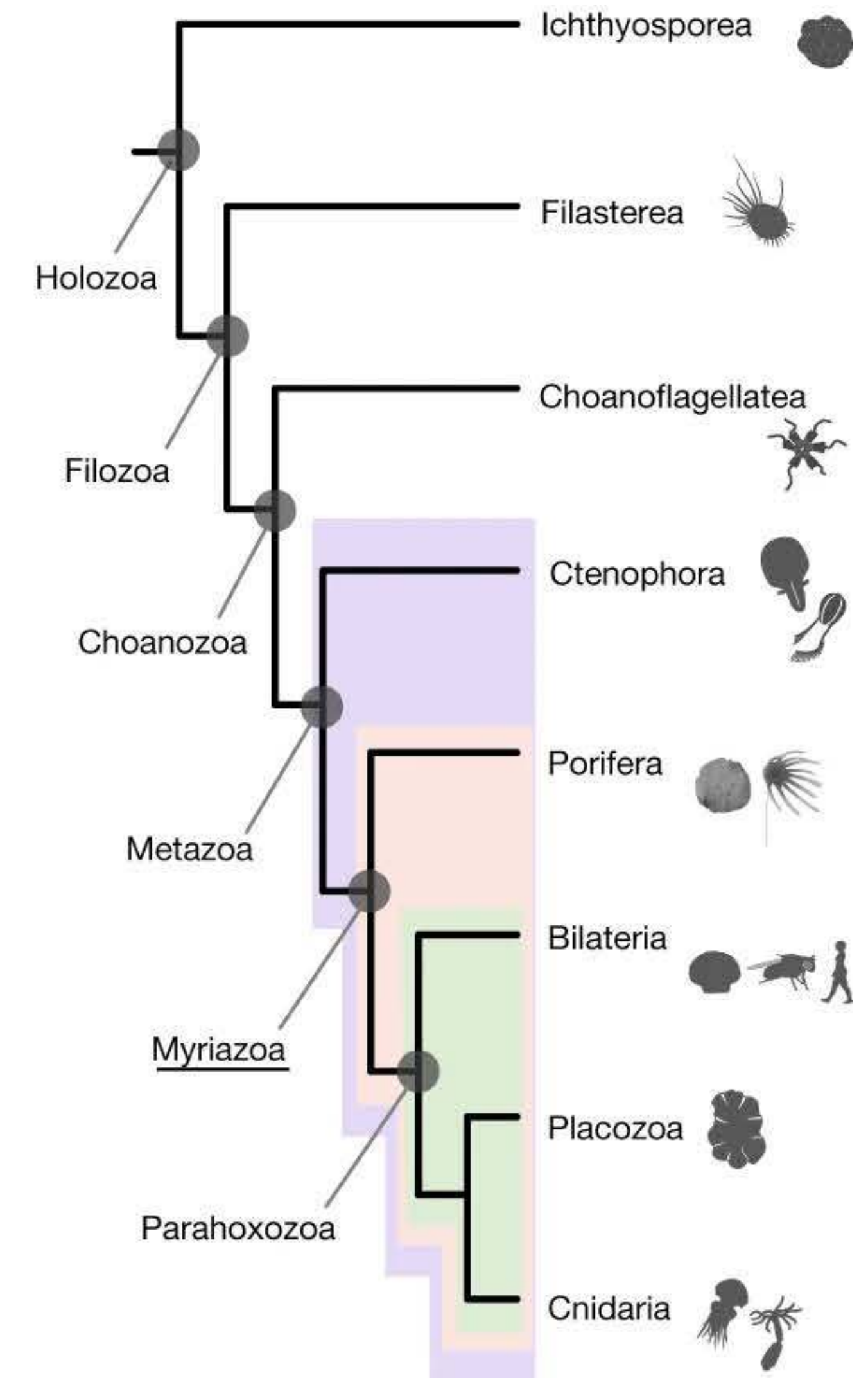


Figure from Schultz et al. ([Nature, 2023](#))

Misconceptions

A jelly is the ultimate ancestor of all animal life, finds study

For nearly 100 years, scientists have thought a sponge was the first type of animal on the planet, but new genetic technique points to a creature like today's comb jellies

Ian Johnston Science Correspondent • Monday 10 April 2017 16:02 BST •



600-800 million years of evolution gave us ctenophores, sponges, jellyfish, and humans...so it is not correct to say “our ancestors were comb jellies”



SMART NEWS

Comb Jellies May Be the World's Oldest Animal Group

A new study suggests that ancestors of comb jellies, not sponges, were the first to break off from the common ancestor of all animals



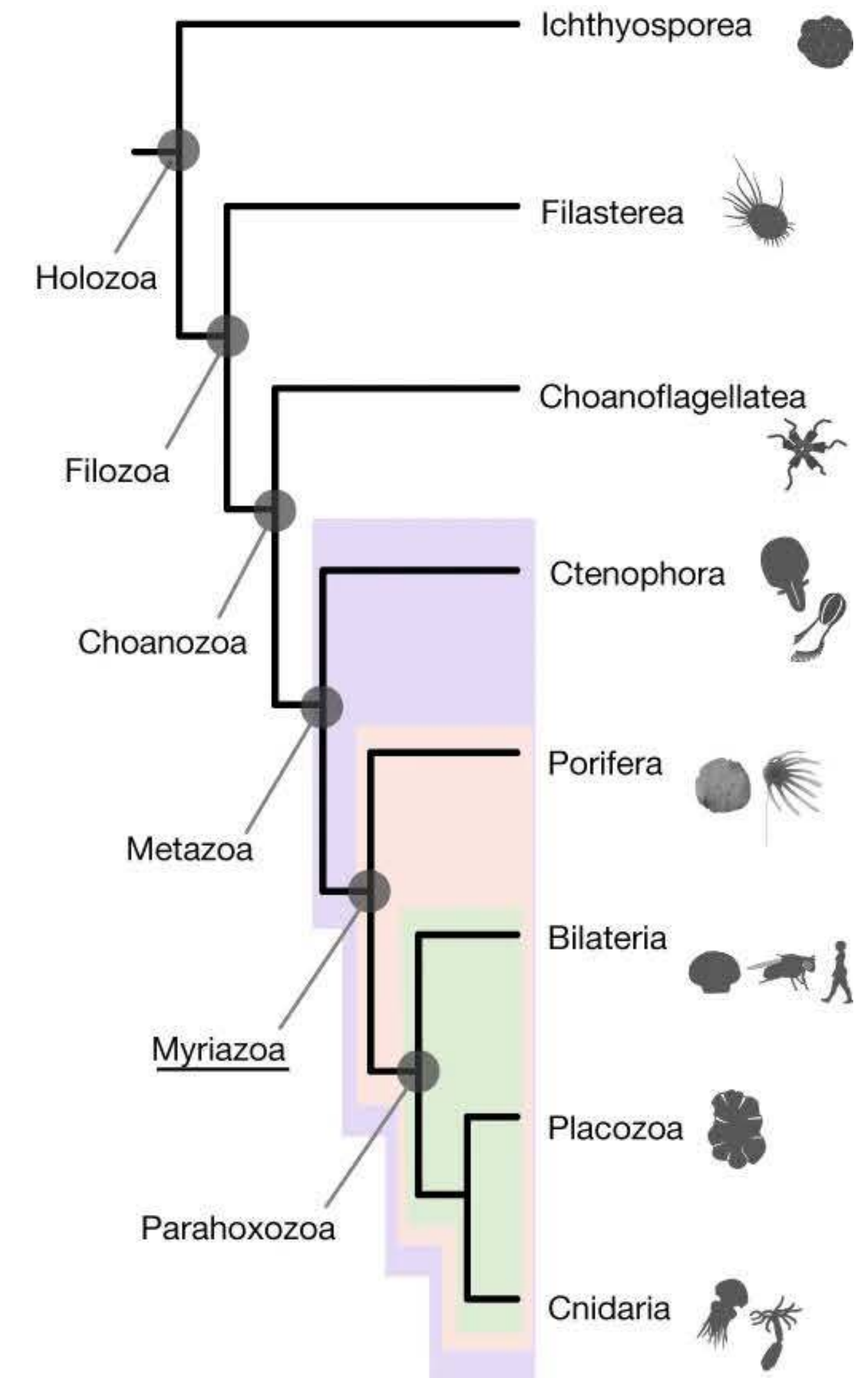
Margaret Osborne

Daily Correspondent

May 18, 2023

Smithsonian
MAGAZINE

The most-recent common ancestor of all known ctenophores is younger than the ancestor of sponges, bilaterians, & cnidarians. A lot has happened during the history of animals that we don't know...

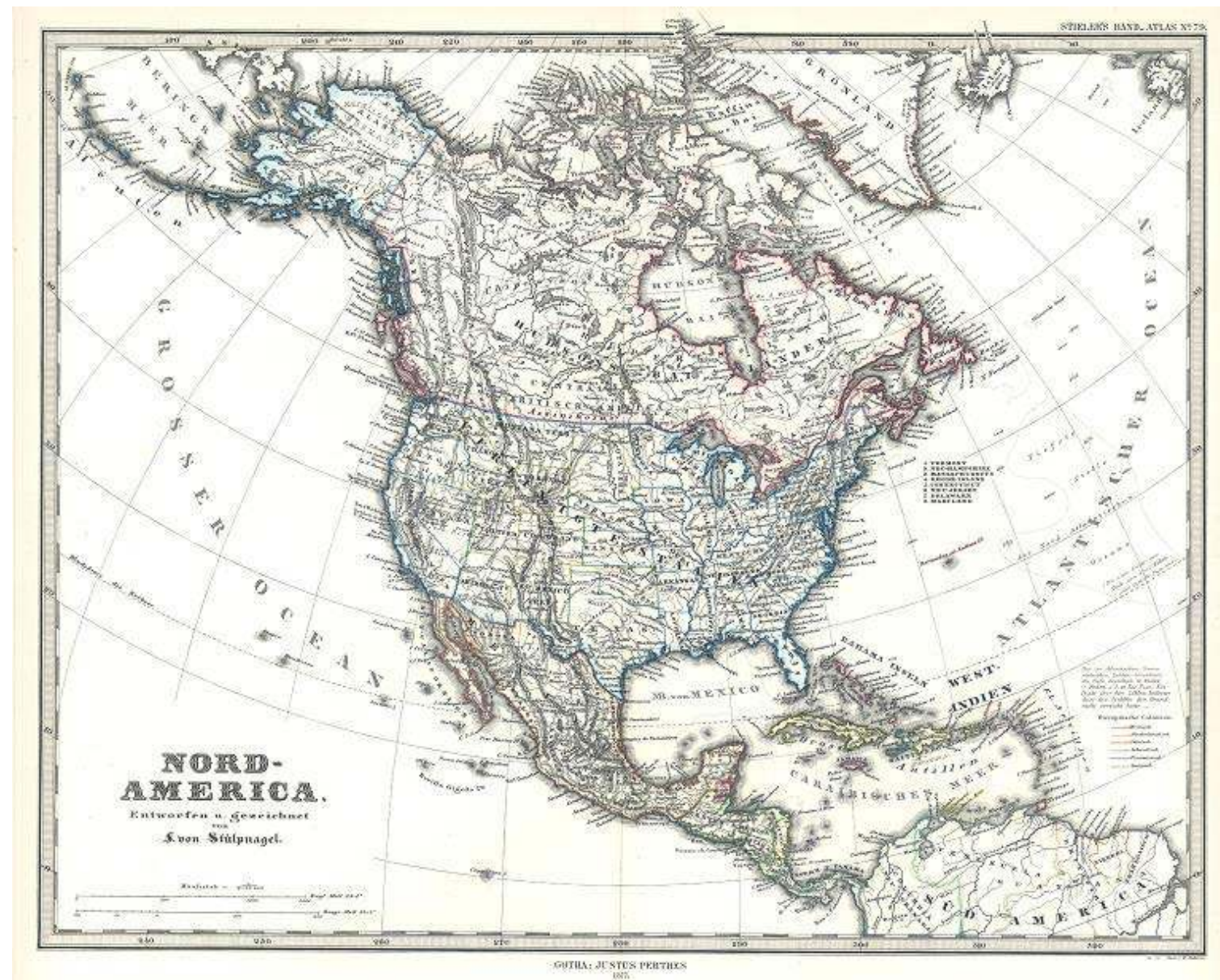


Estimates Change with More Data

1780



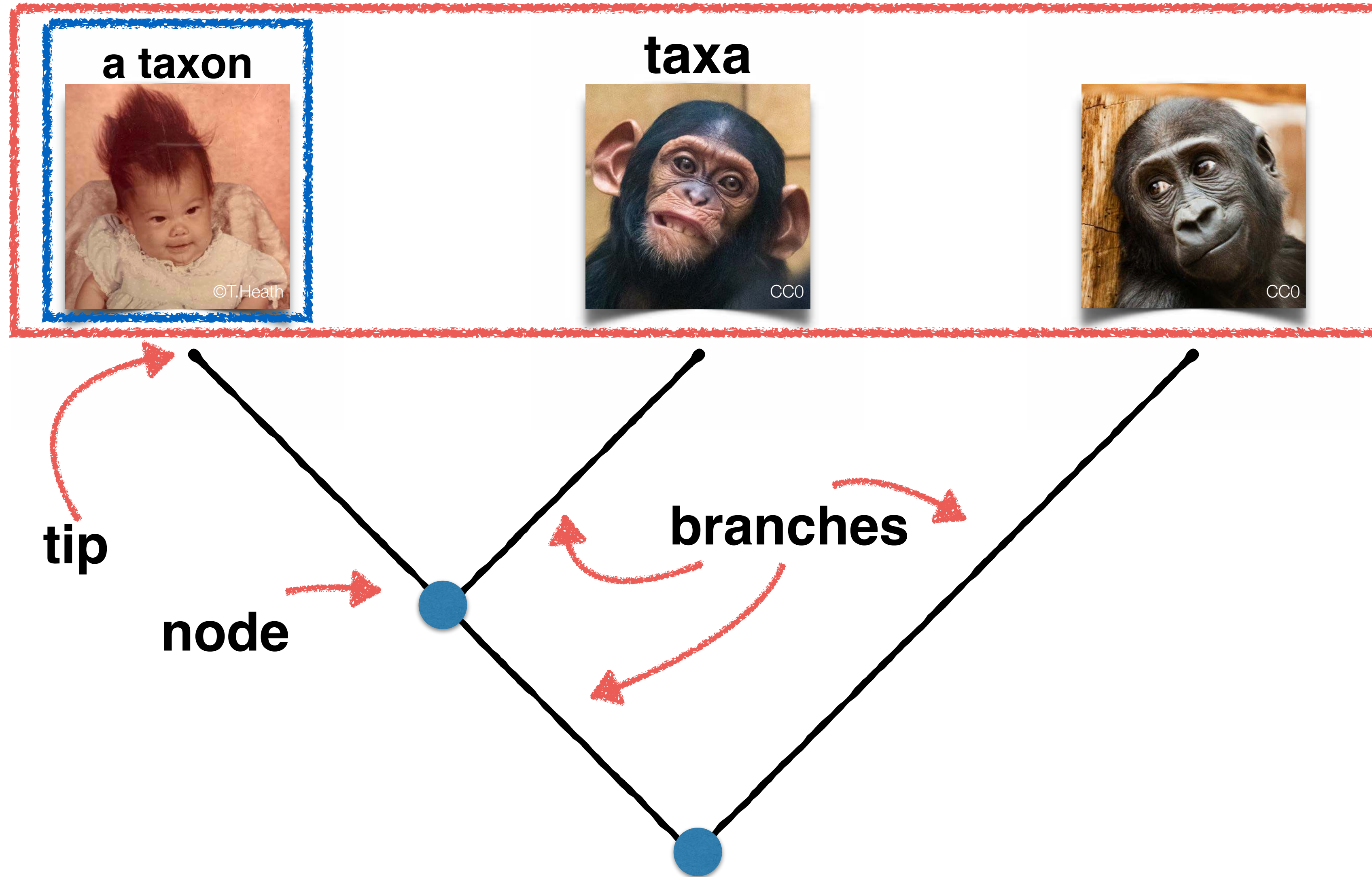
1877



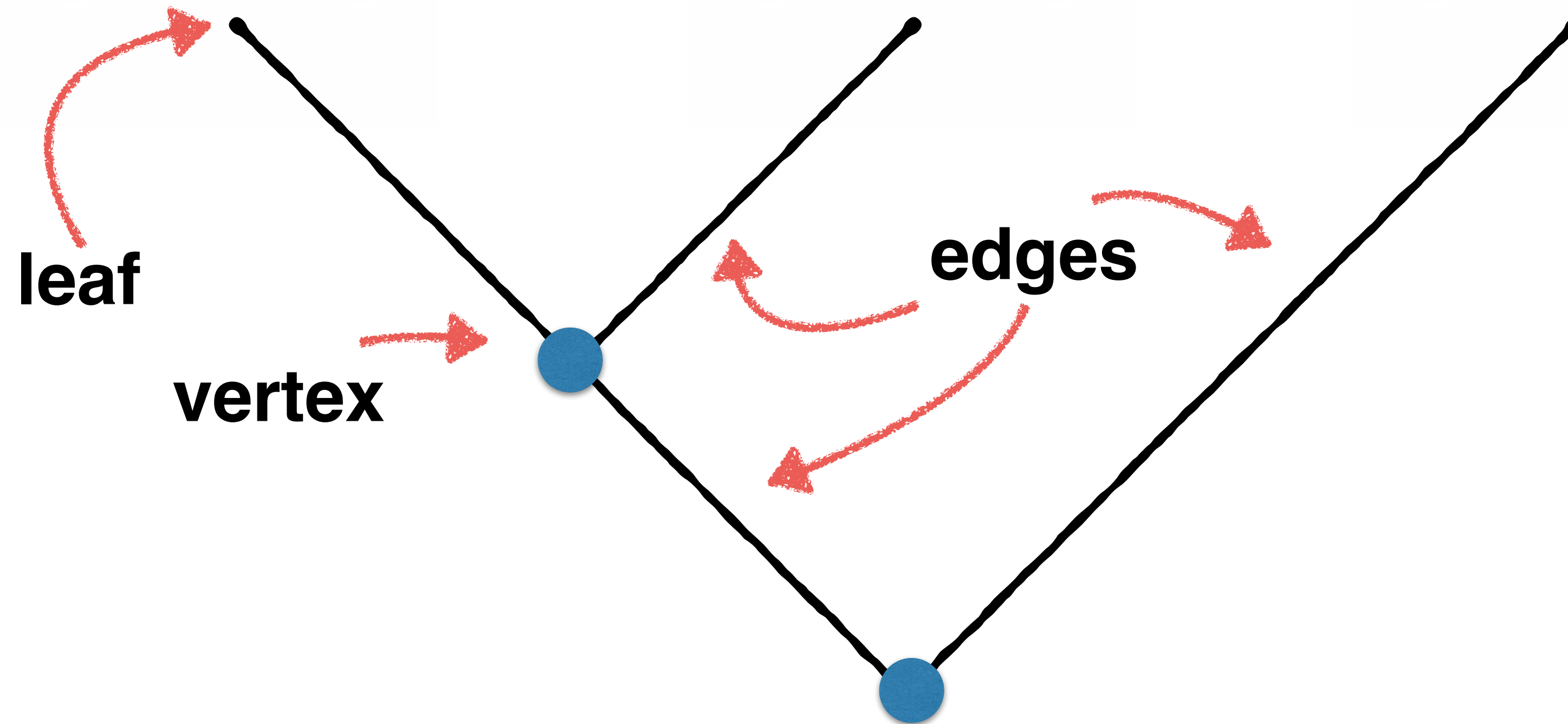
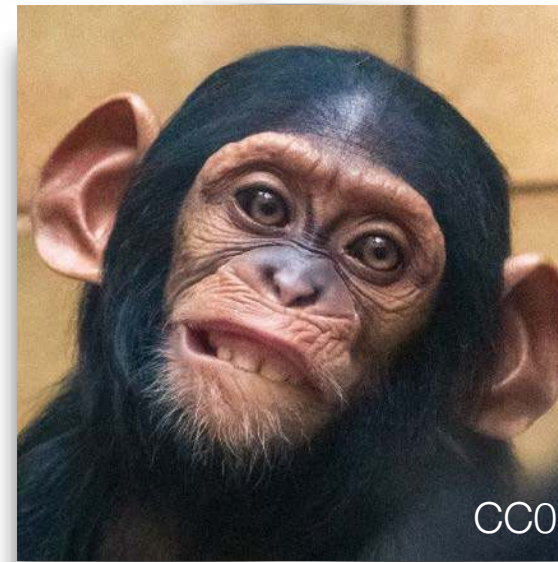
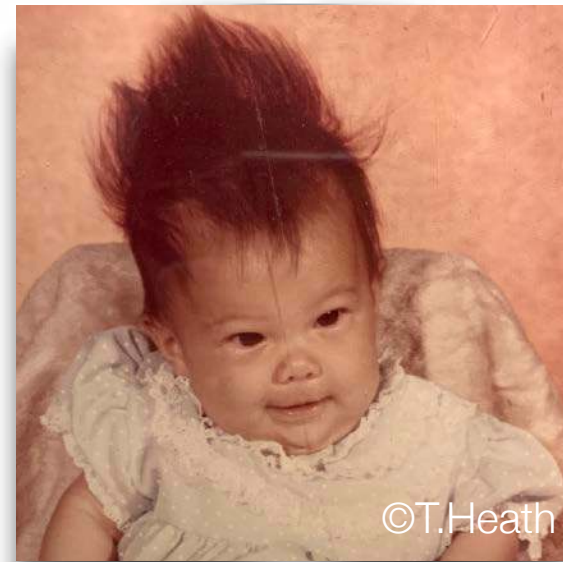
Like maps, phylogenetic trees are **estimates** that can change (and improve) with **more data** and **better methods**

The Metazoa phylogeny may change again...

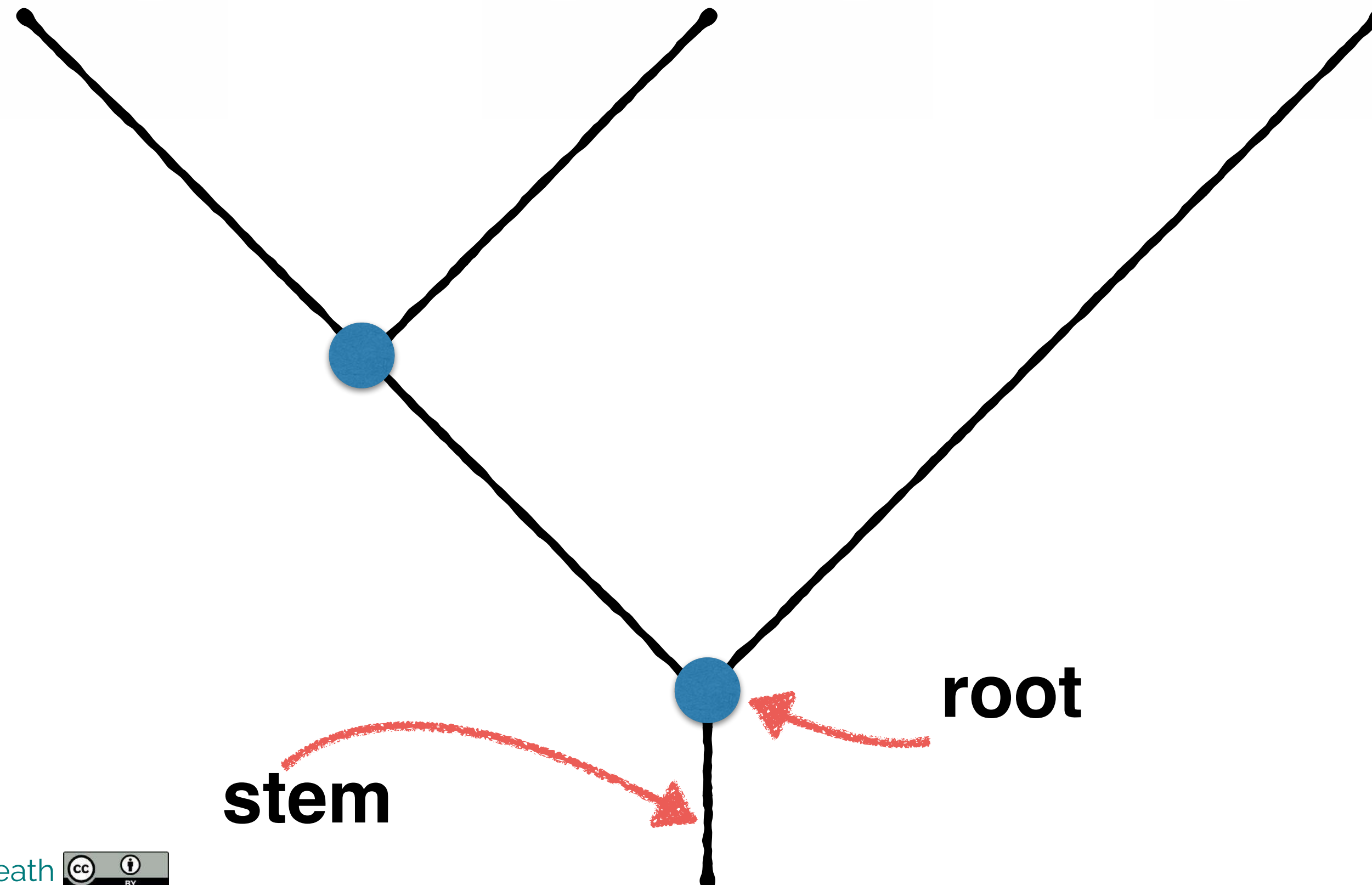
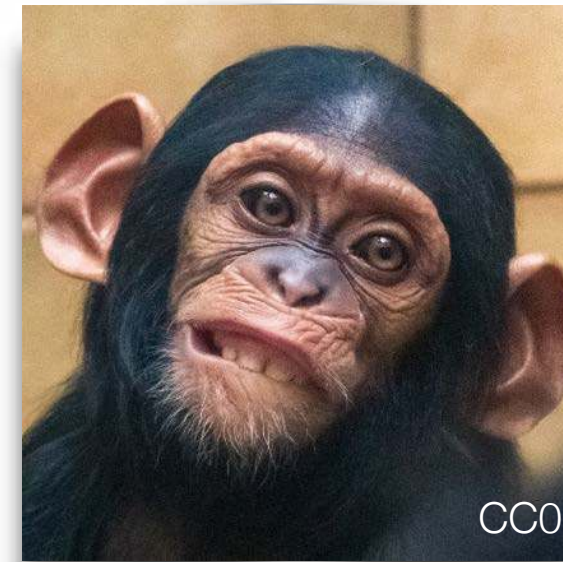
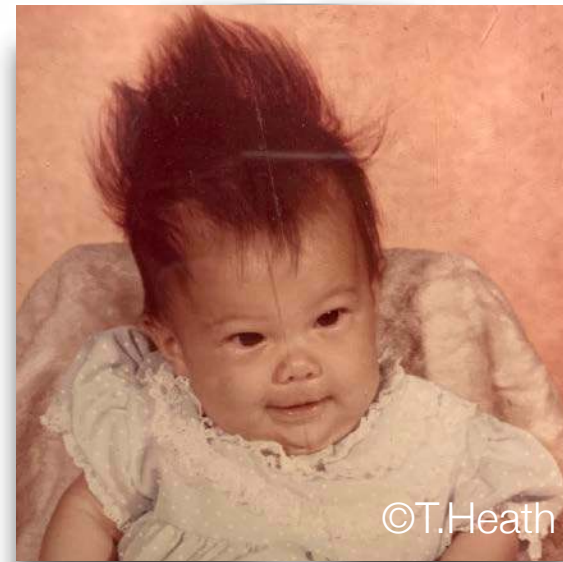
Phylogeny Terminology & Concepts



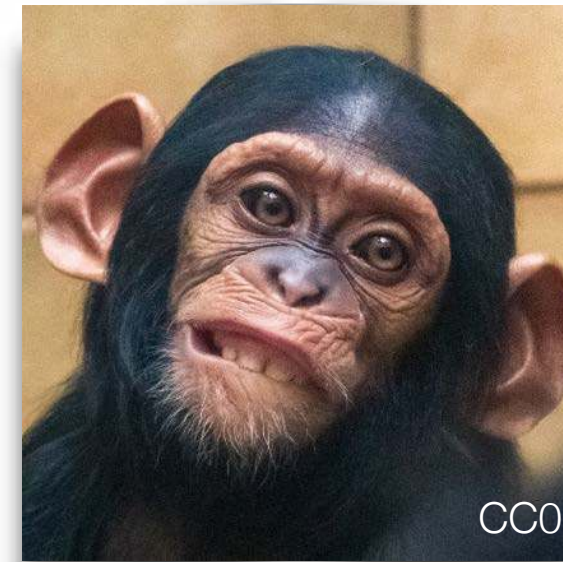
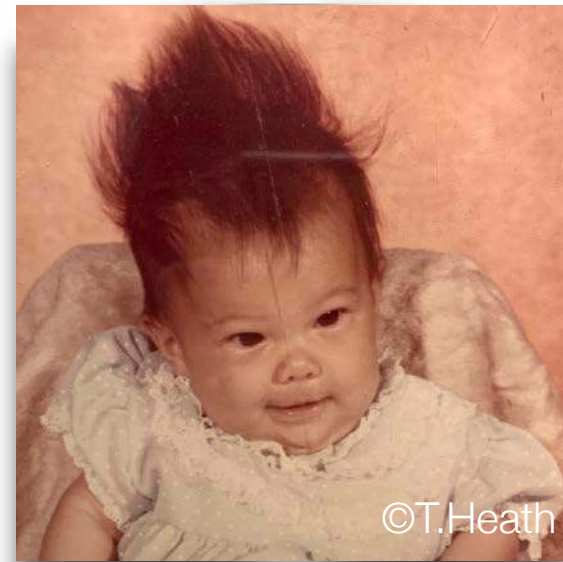
Phylogeny Terminology & Concepts



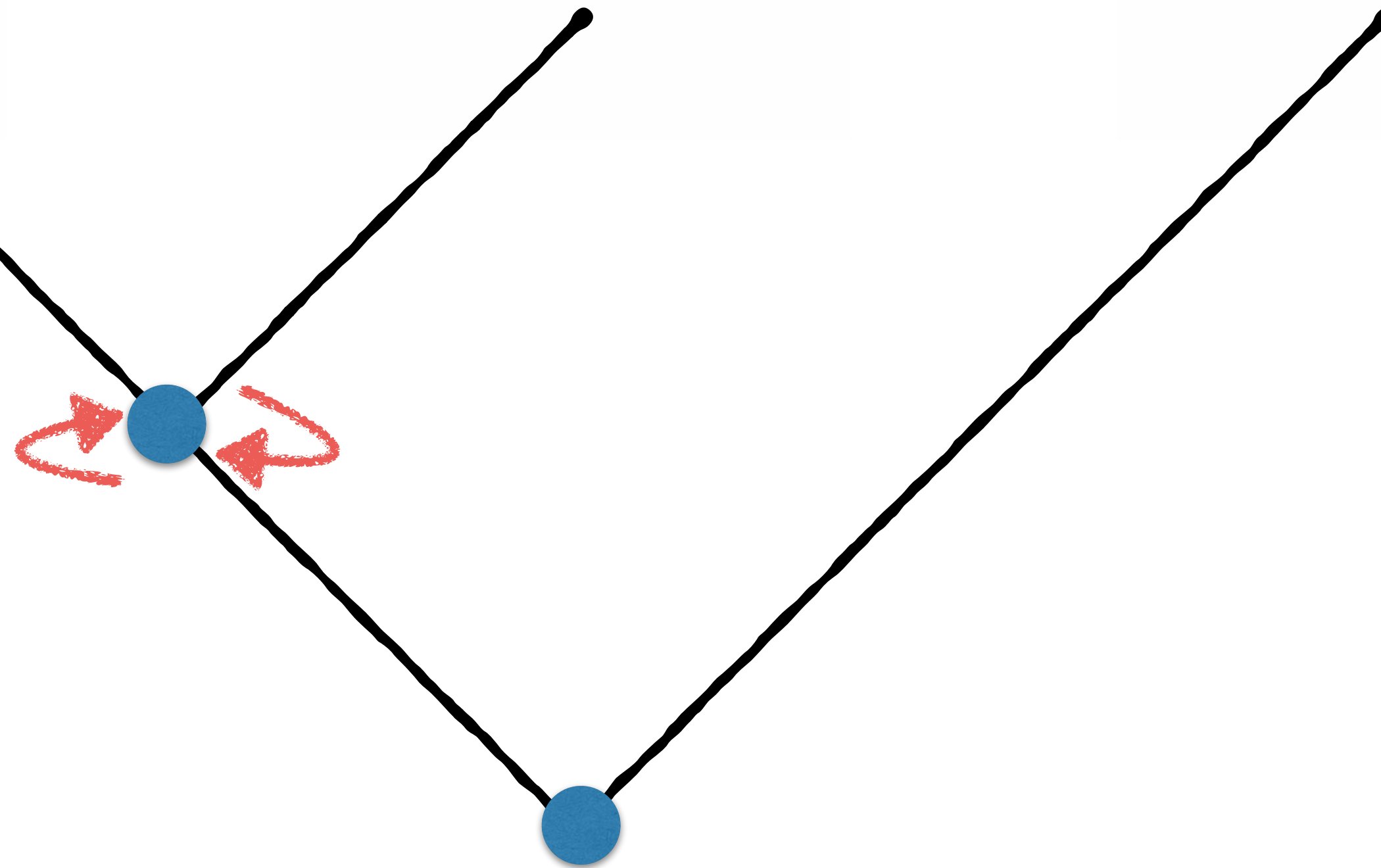
Phylogeny Terminology & Concepts



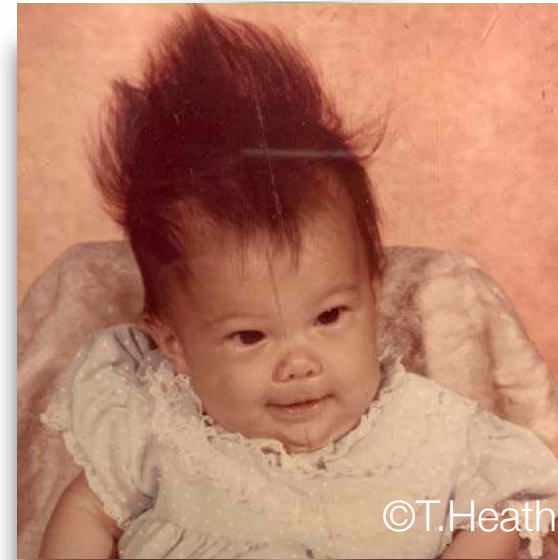
Phylogeny Terminology & Concepts



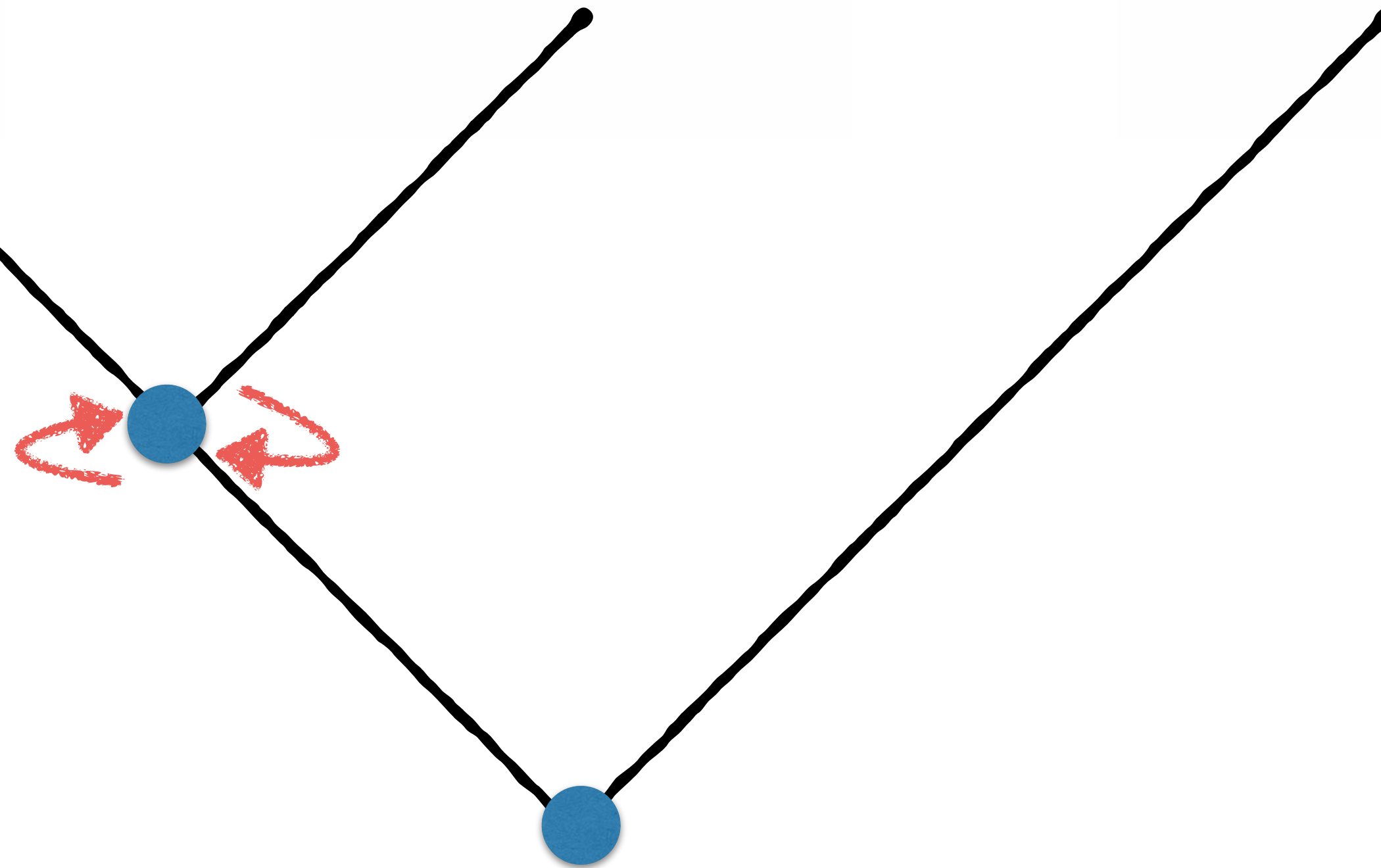
nodes can be rotated
without changing the
relationships of the
descendant branches



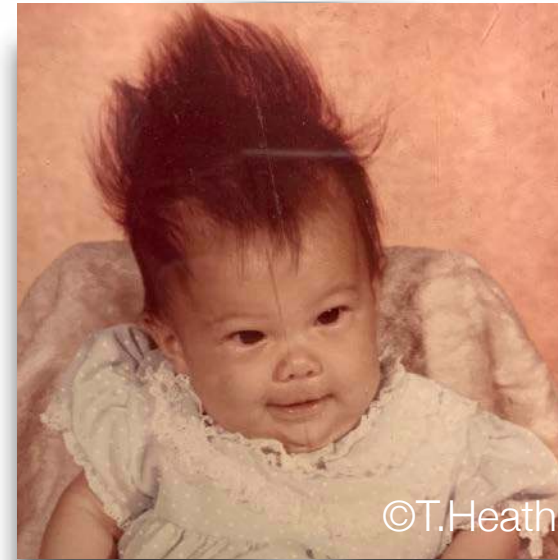
Phylogeny Terminology & Concepts



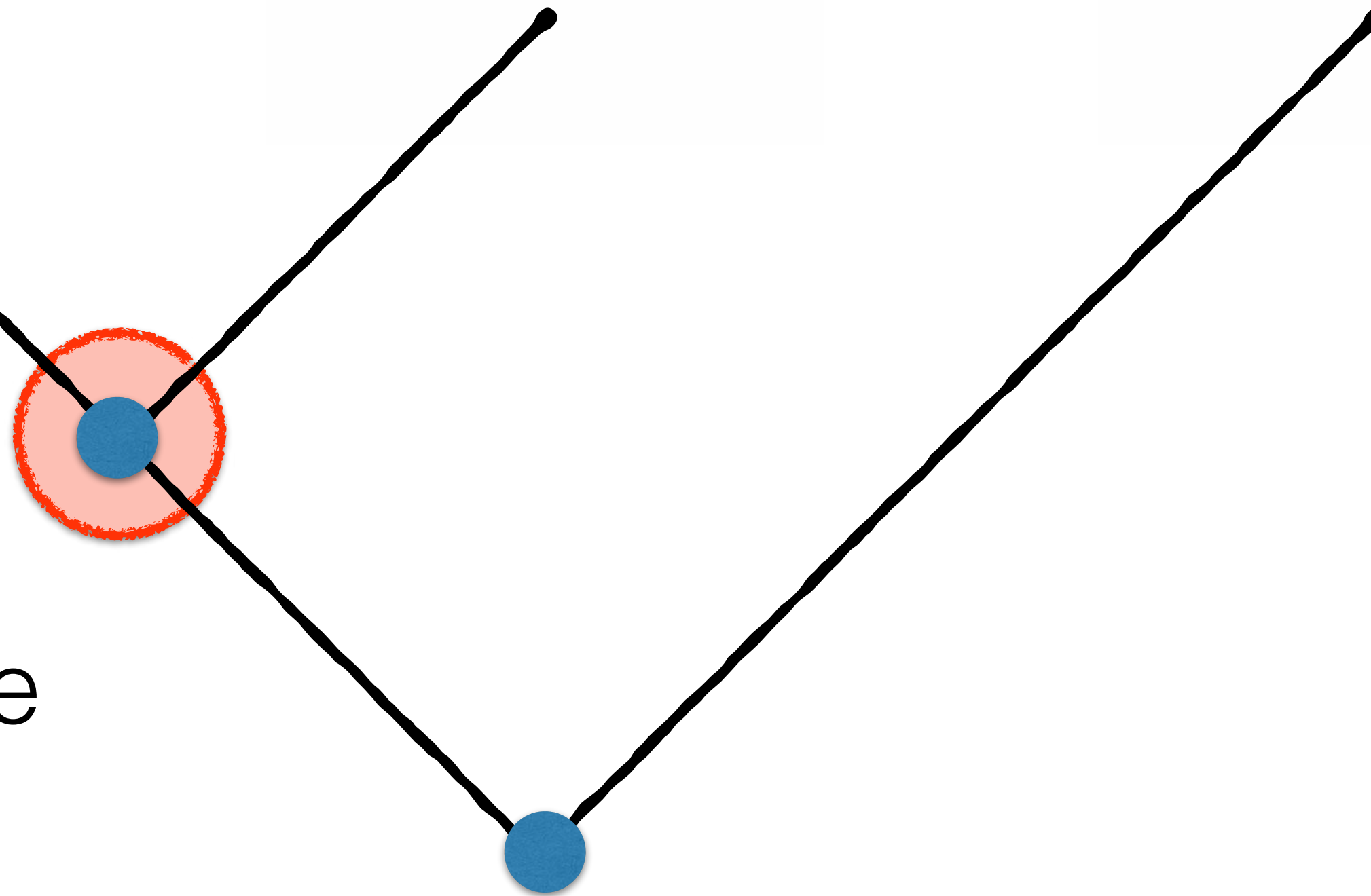
nodes can be rotated
without changing the
relationships of the
descendant branches



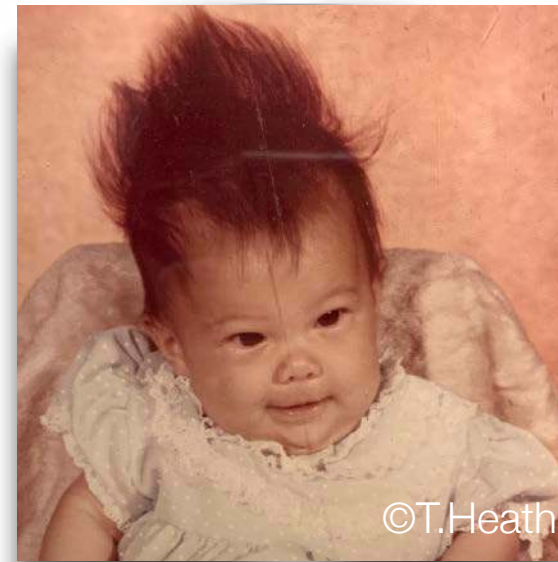
Phylogeny Terminology & Concepts



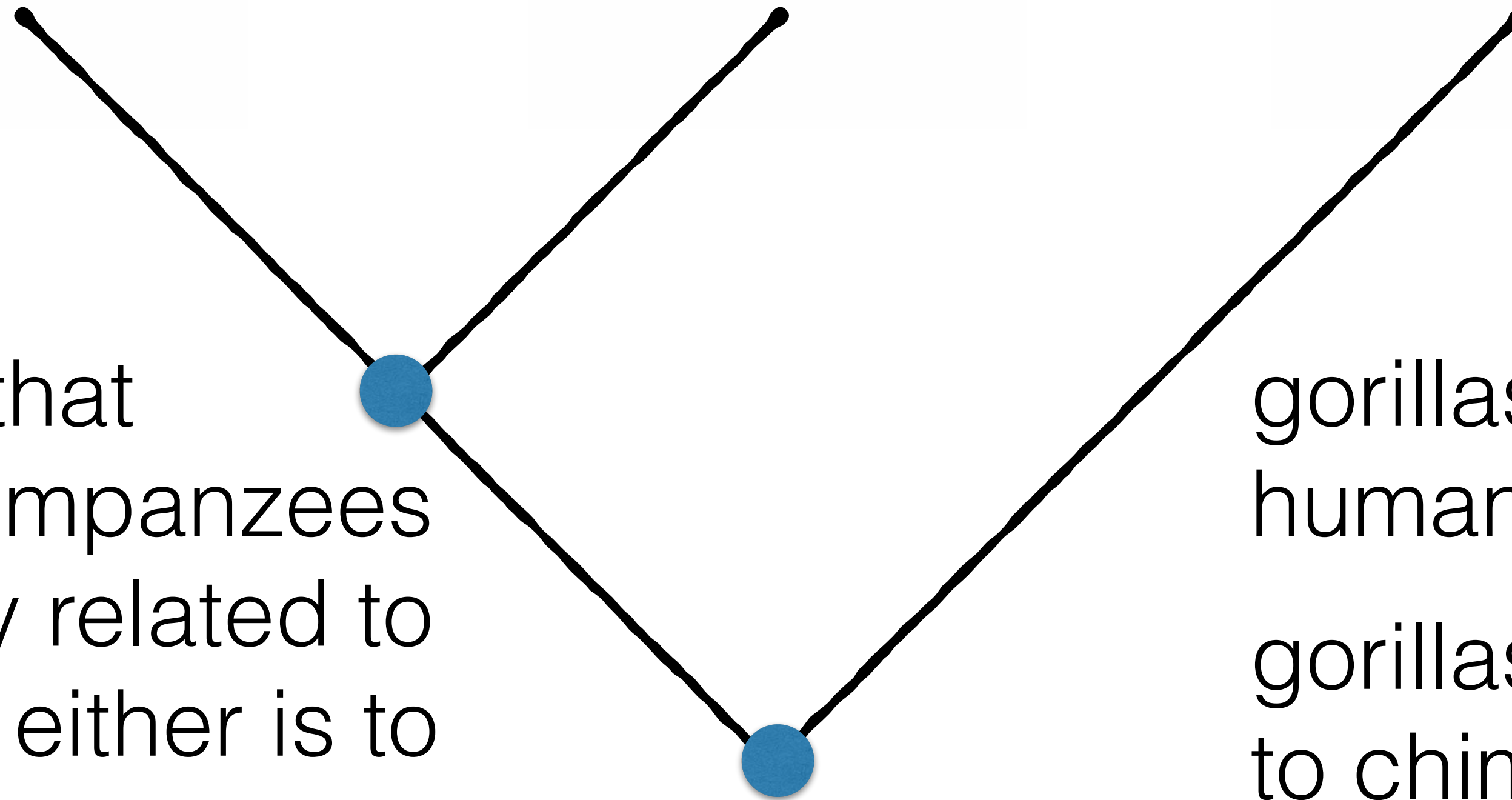
a node represents the
**most-recent common
ancestor** (MRCA) of the
descendant taxa



Phylogeny Terminology & Concepts

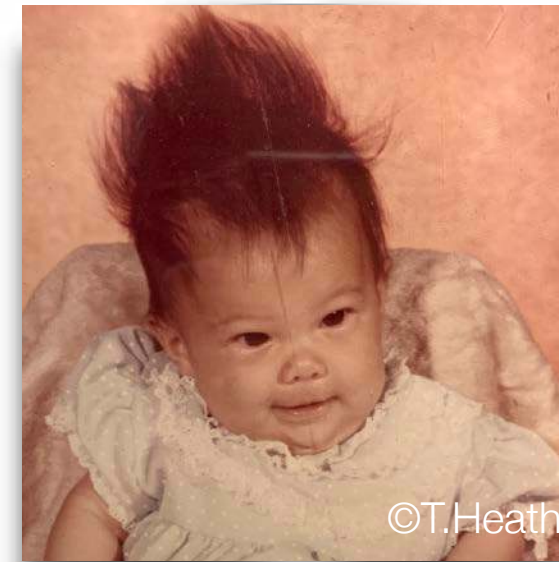


this tree shows that
humans and chimpanzees
are more closely related to
each other than either is to
gorillas



gorillas *are sister to*
humans & chimps
gorillas *are NOT basal*
to chimps & humans

Phylogeny Terminology & Concepts

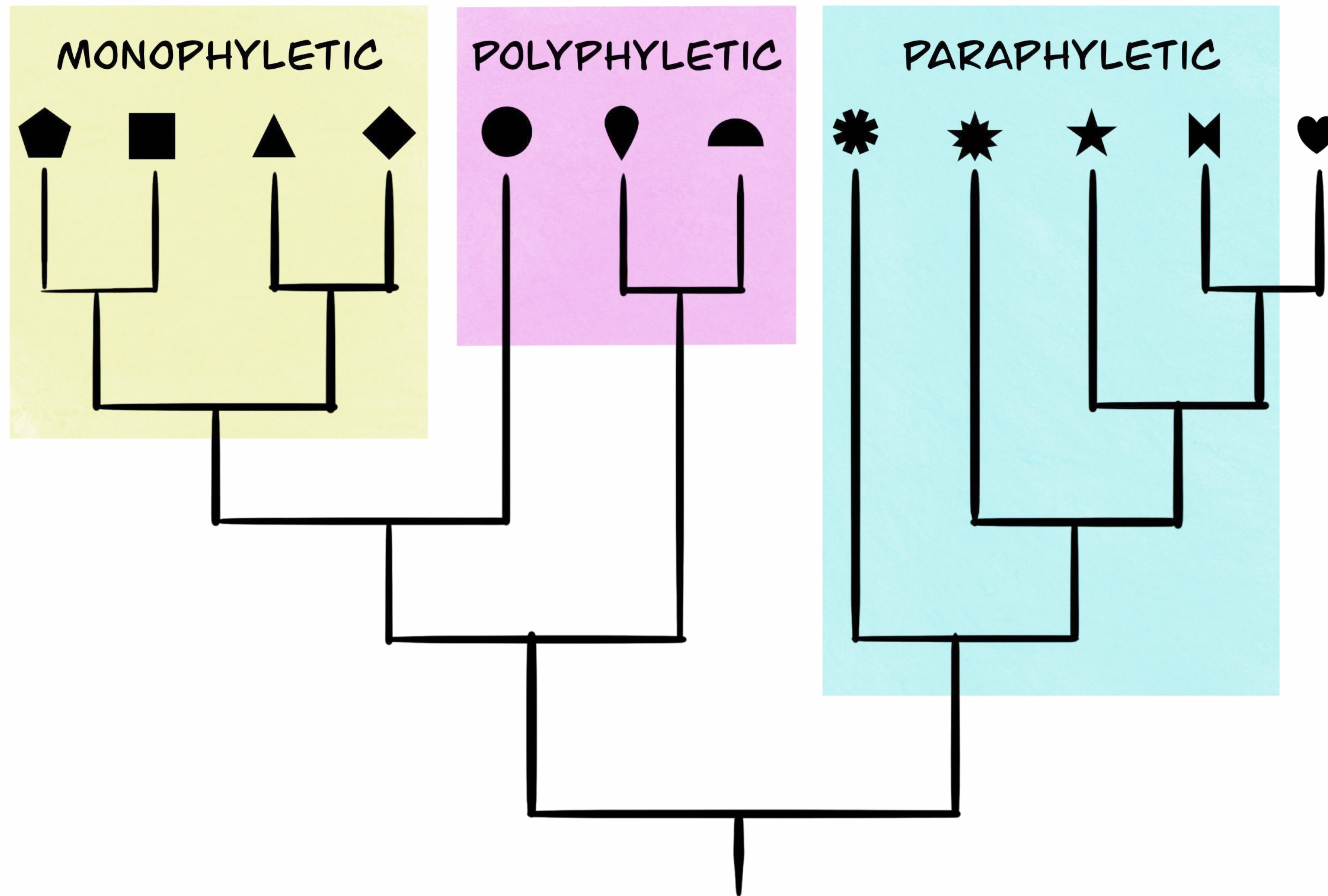


the term basal refers to something that is "closer to the base", so please never use this for extant taxa



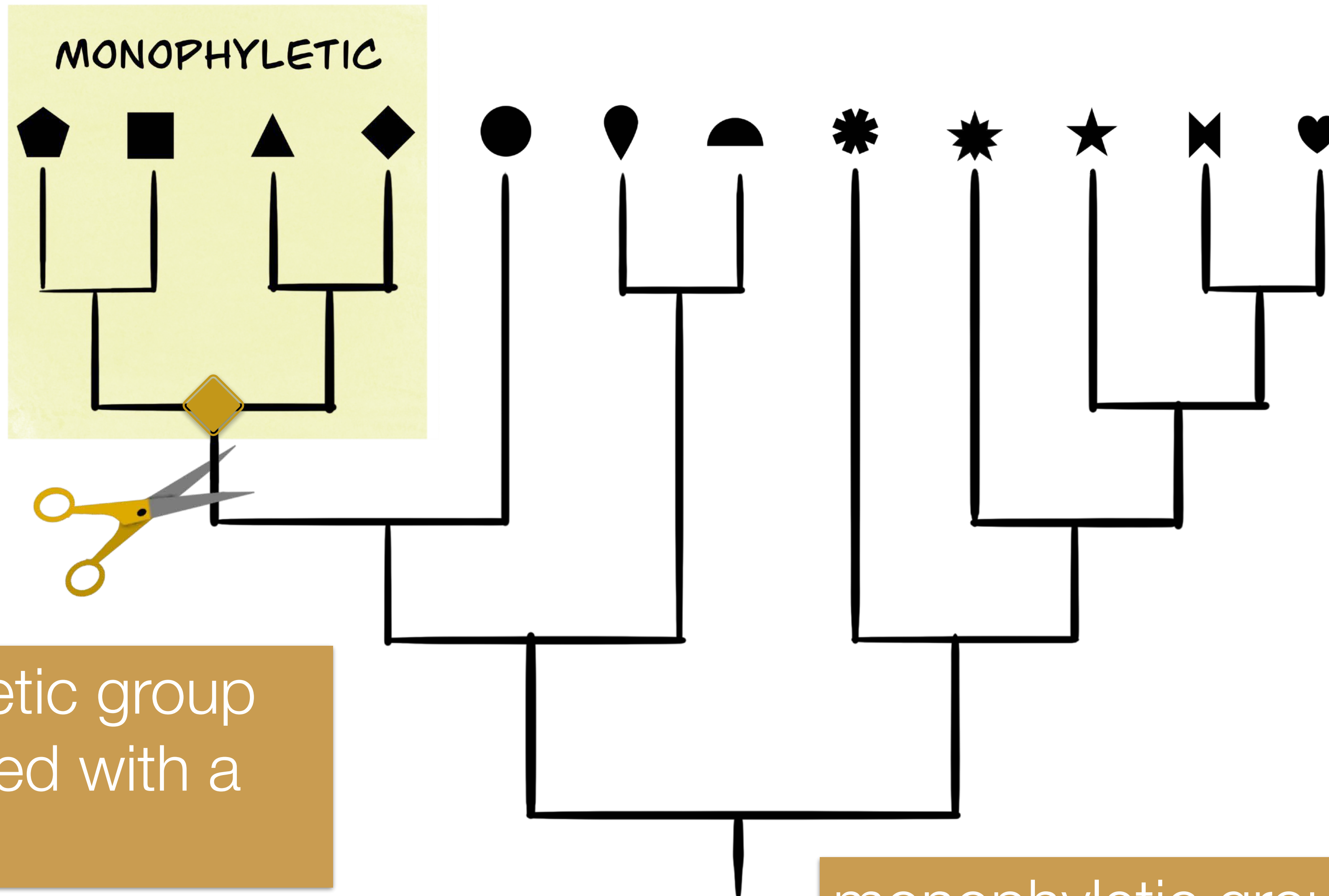
the 10 Mya fossil *Chororapithecus abyssinicus* can be called a basal ape
(Sewa et al 2007)

Groups in a Phylogeny



Monophyly

a monophyletic group includes an ancestor and all of its descendants

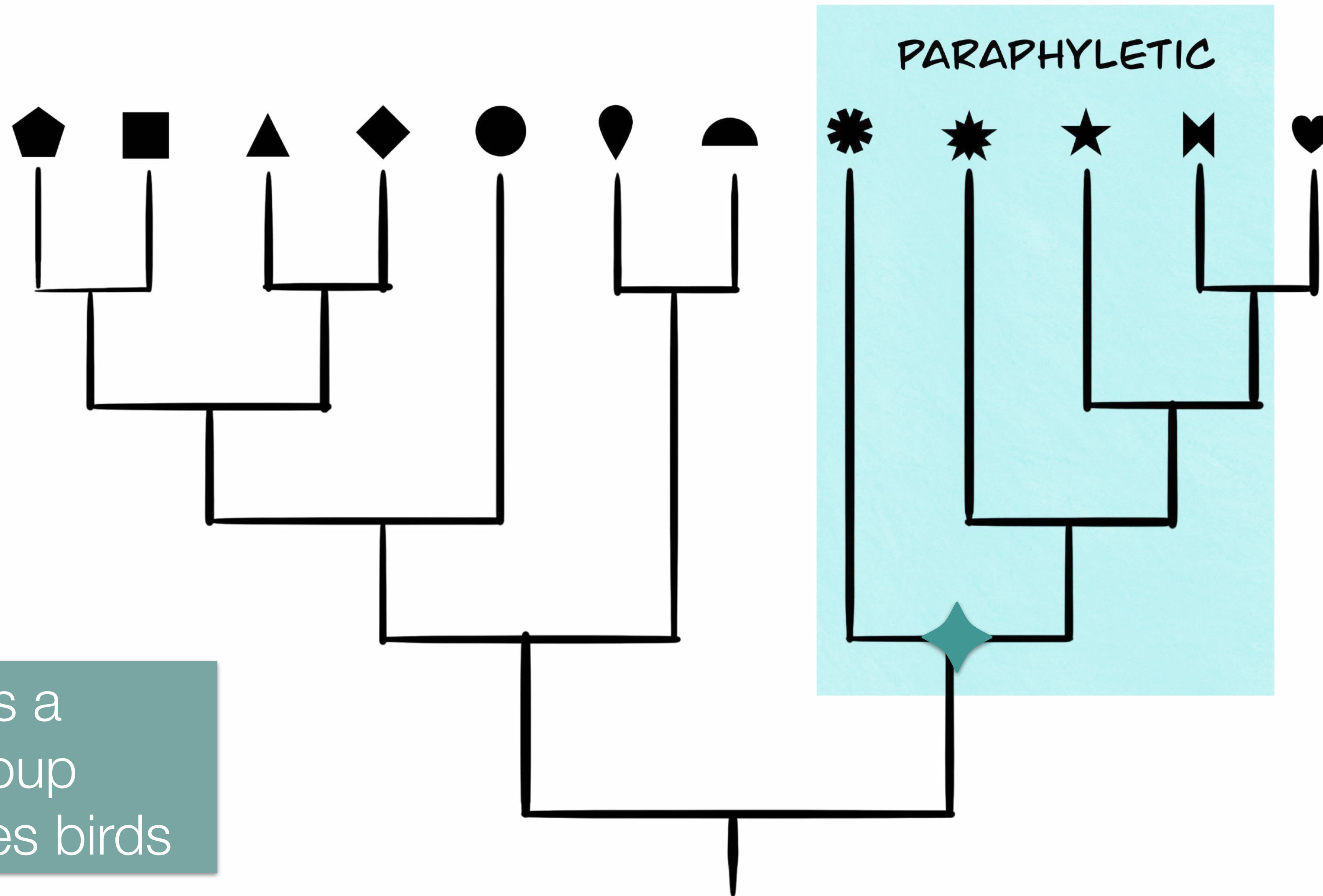


a monophyletic group can be pruned with a single cut

monophyletic group = clade

Paraphyly

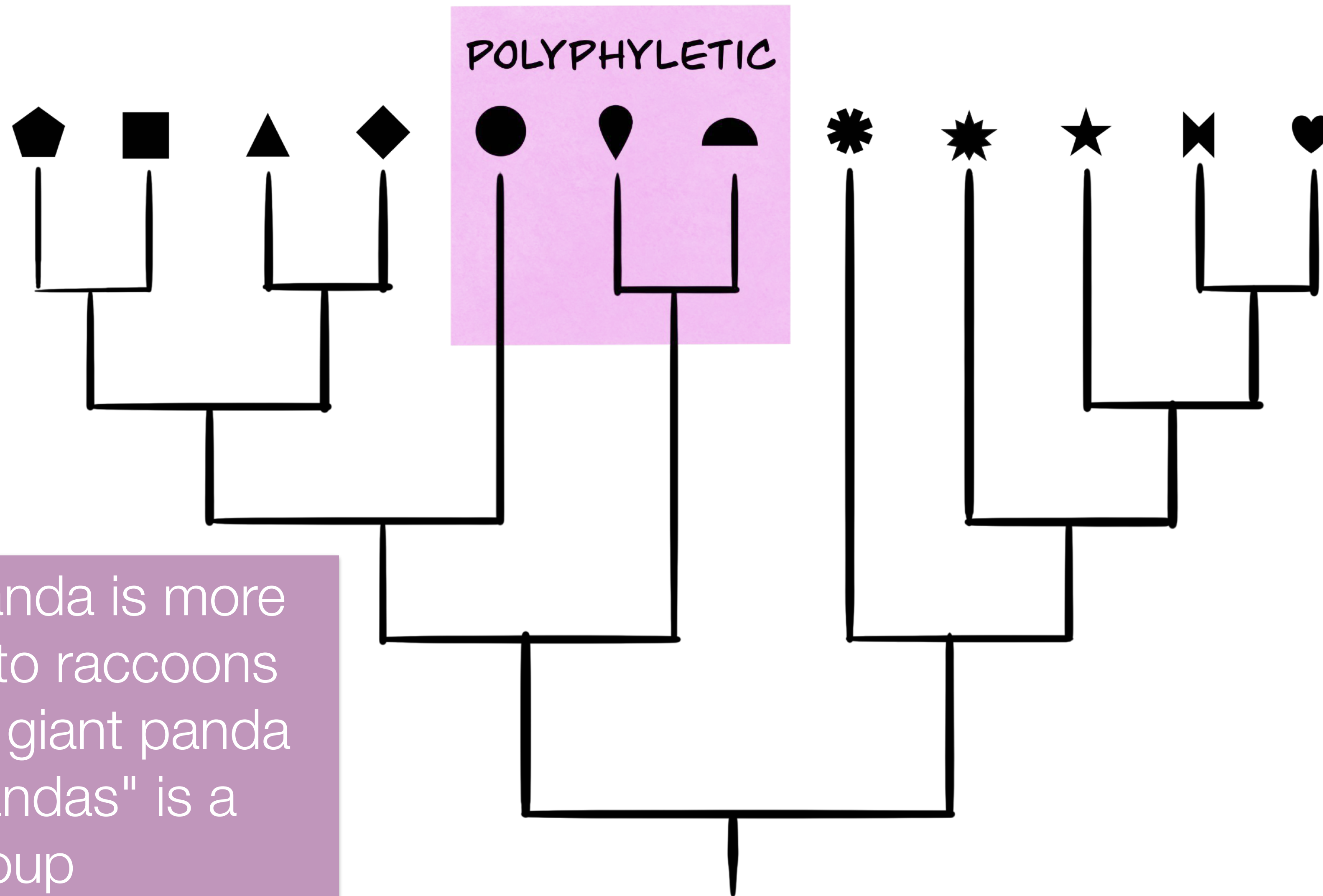
a paraphyletic group includes an ancestor and a subset of its descendants



e.g., "reptiles" is a paraphyletic group unless it includes birds

Polyphyly

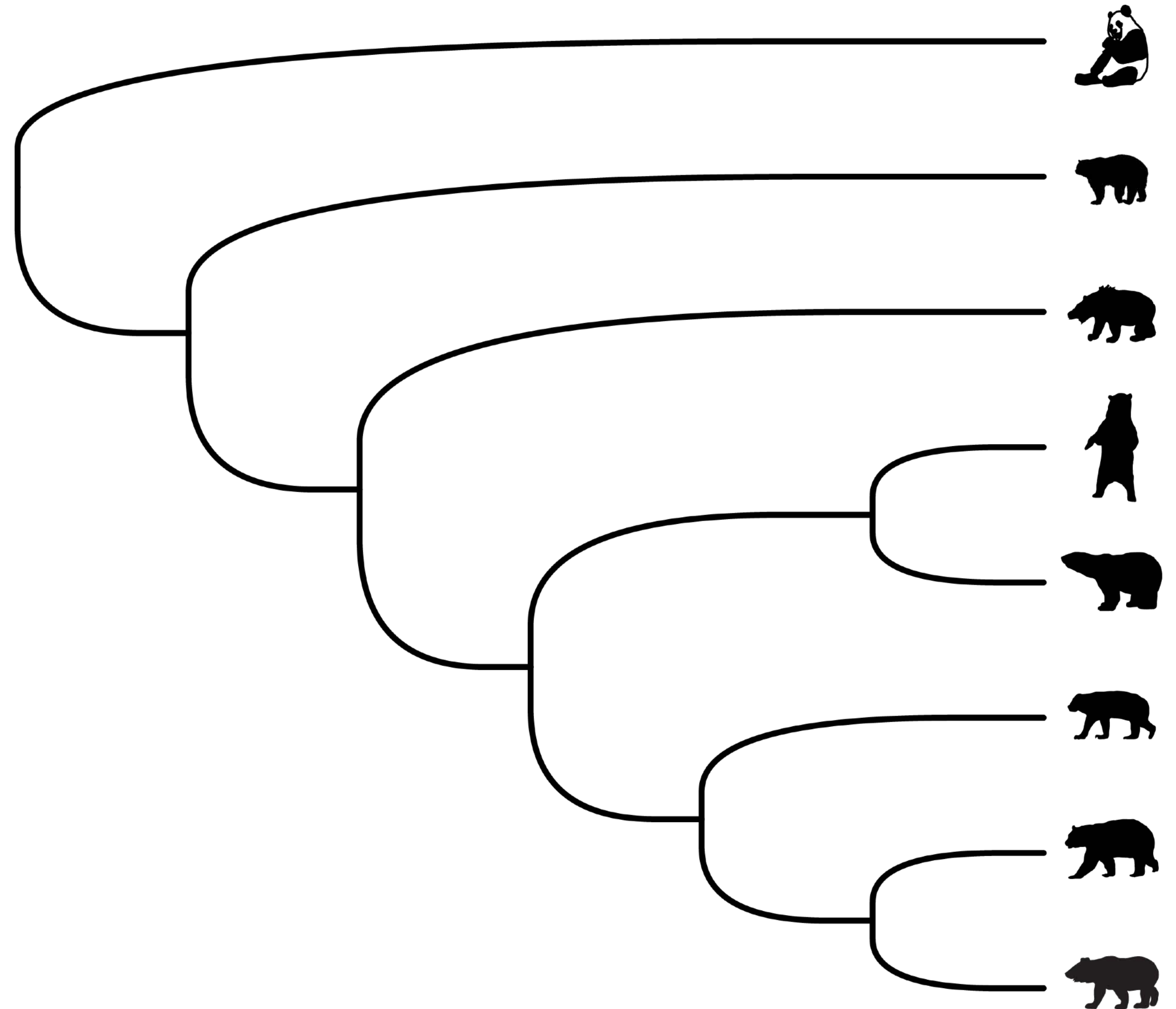
a polyphyletic group includes a set of taxa, but not their common ancestor



e.g., the red panda is more closely related to raccoons than it is to the giant panda (a bear), so "pandas" is a polyphyletic group

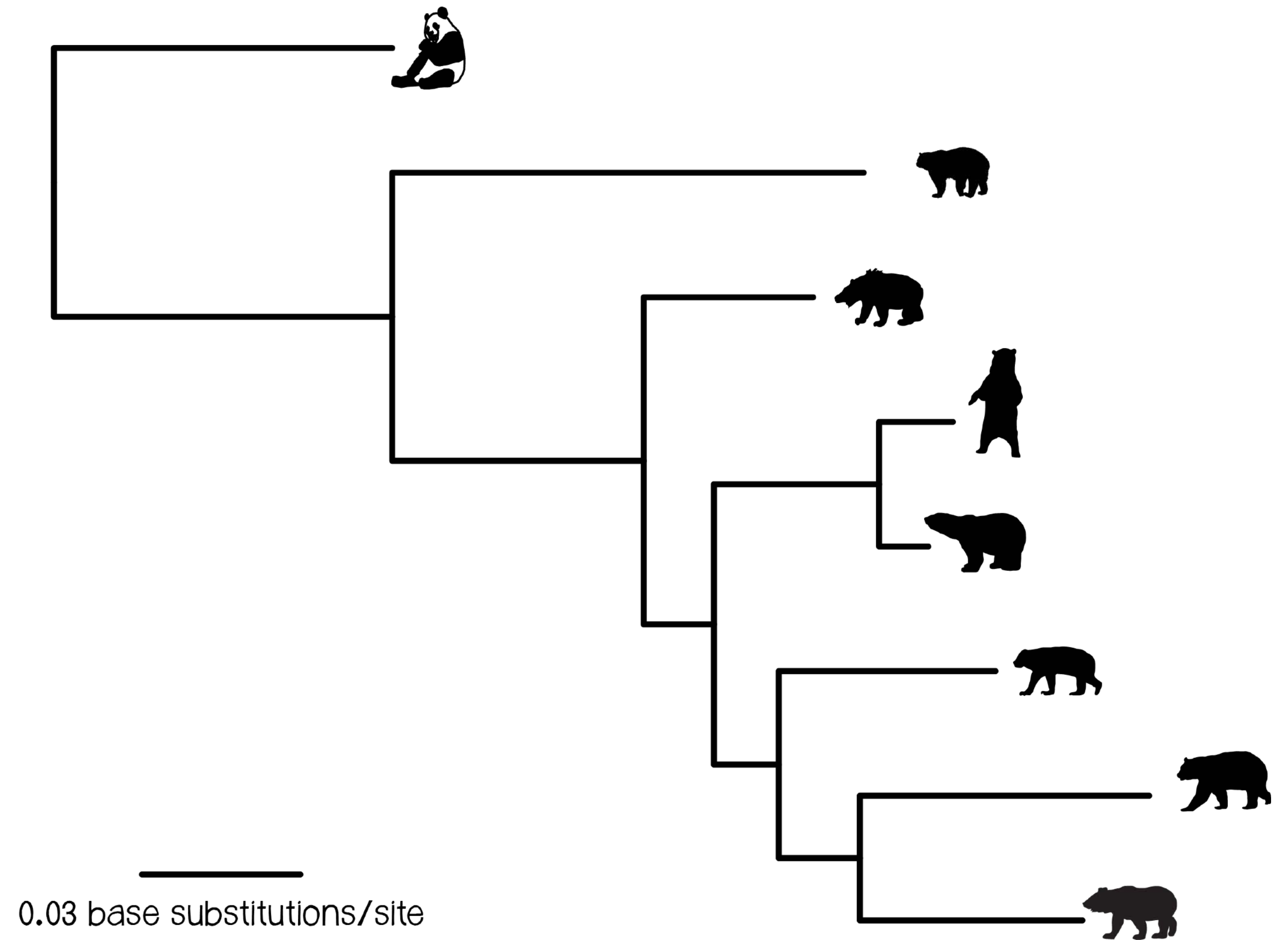
Phylogeny: Branch Lengths

can have no meaning
& just show the pattern
of relationships



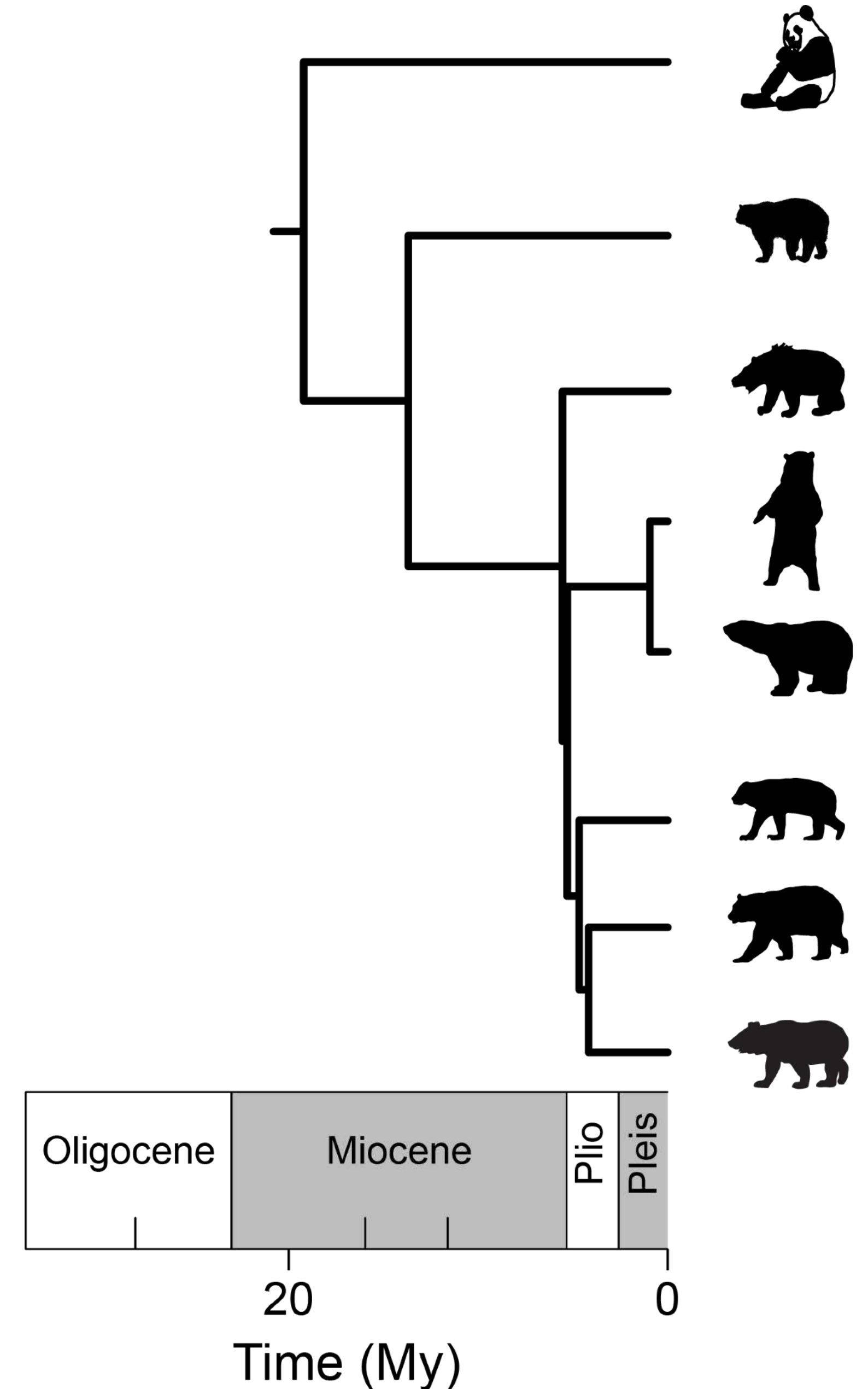
Phylogeny: Branch Lengths

can represent the
amount of **genetic
difference**



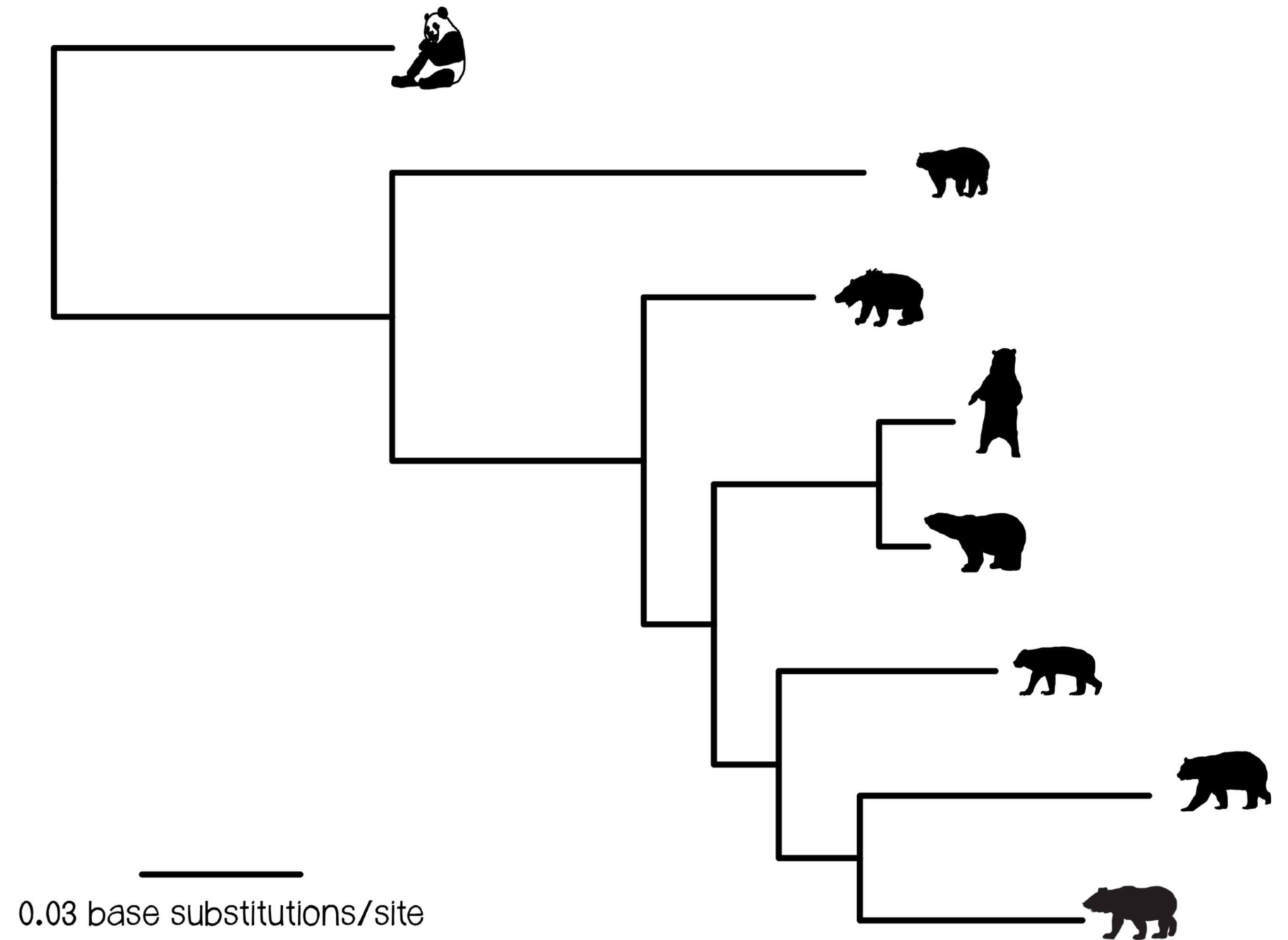
Phylogeny: Branch Lengths

can represent the **duration of time**
between nodes



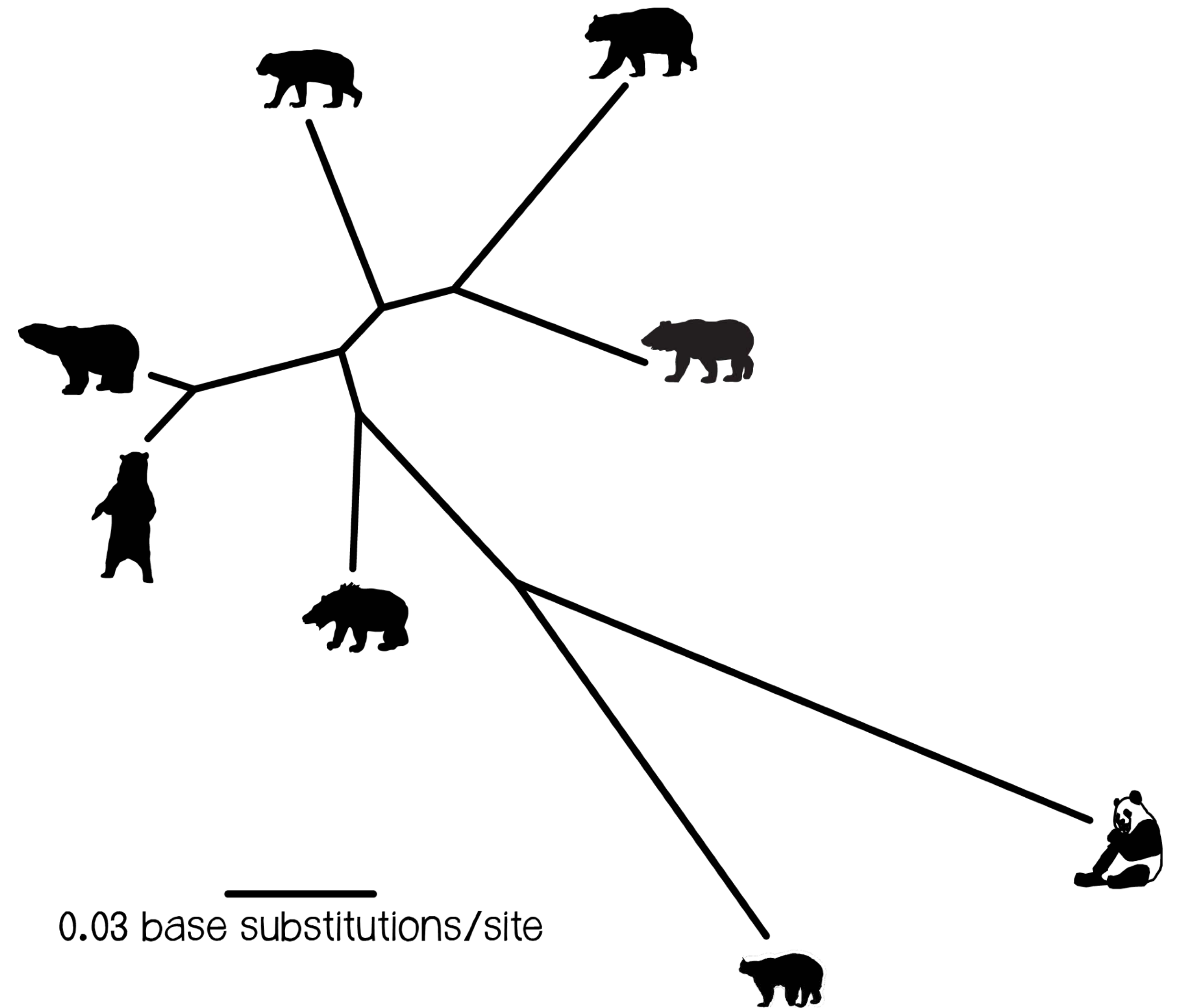
Phylogeny: Rooting

a tree can be **rooted**
to show the direction
or relative timing of
divergence



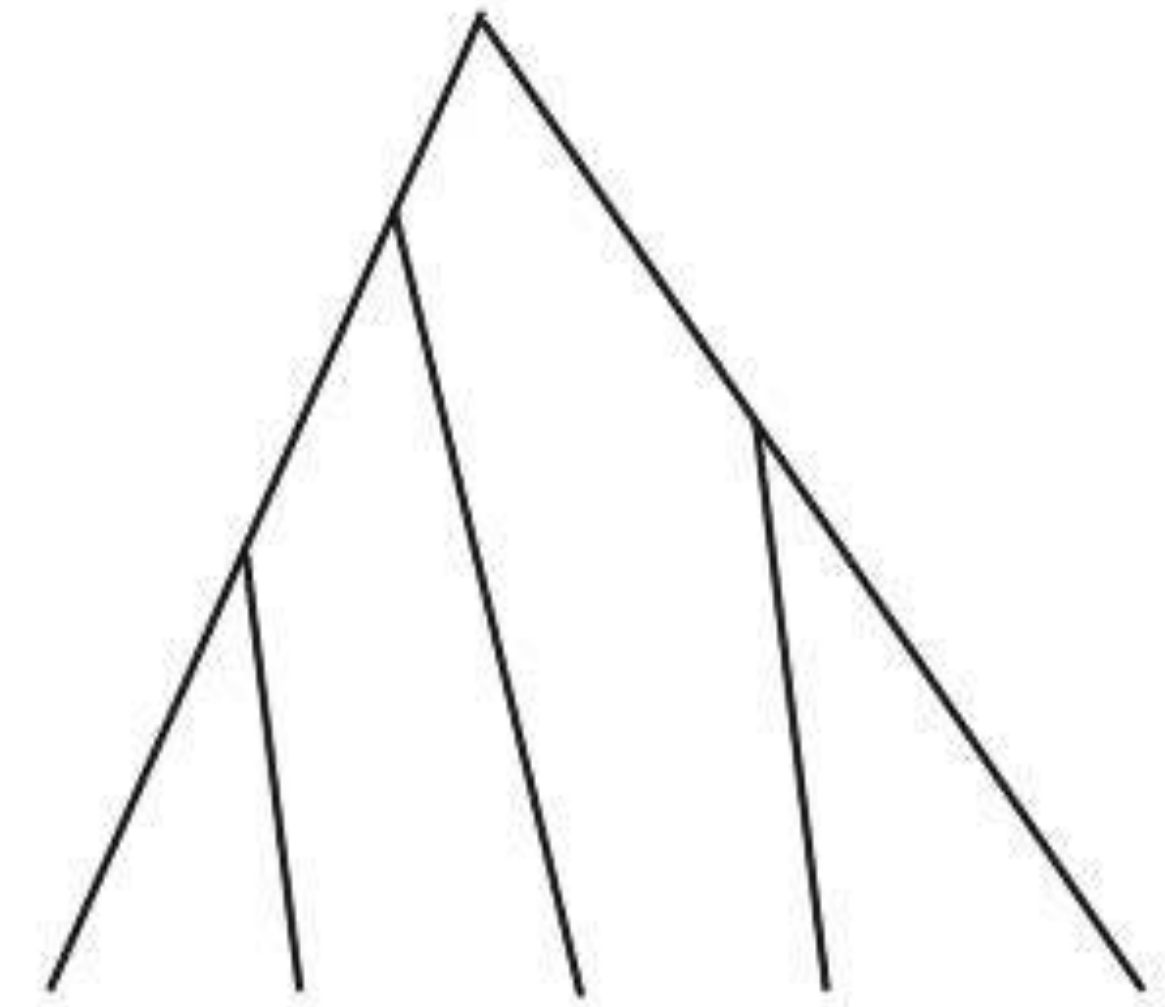
Phylogeny: Rooting

a tree can be
unrooted showing
only relationships
among lineages

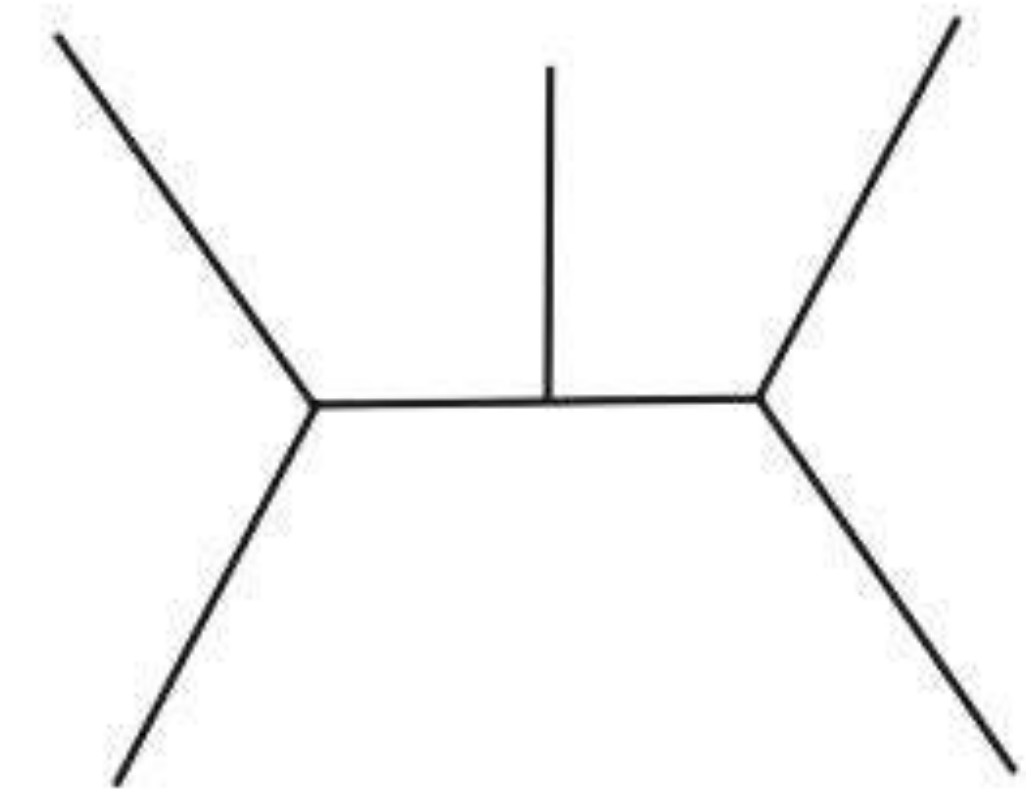


Resolved Phylogeny

a tree can be **bifurcating** or **binary** when all nodes split into only 2 descendants (such a tree is also called "resolved")

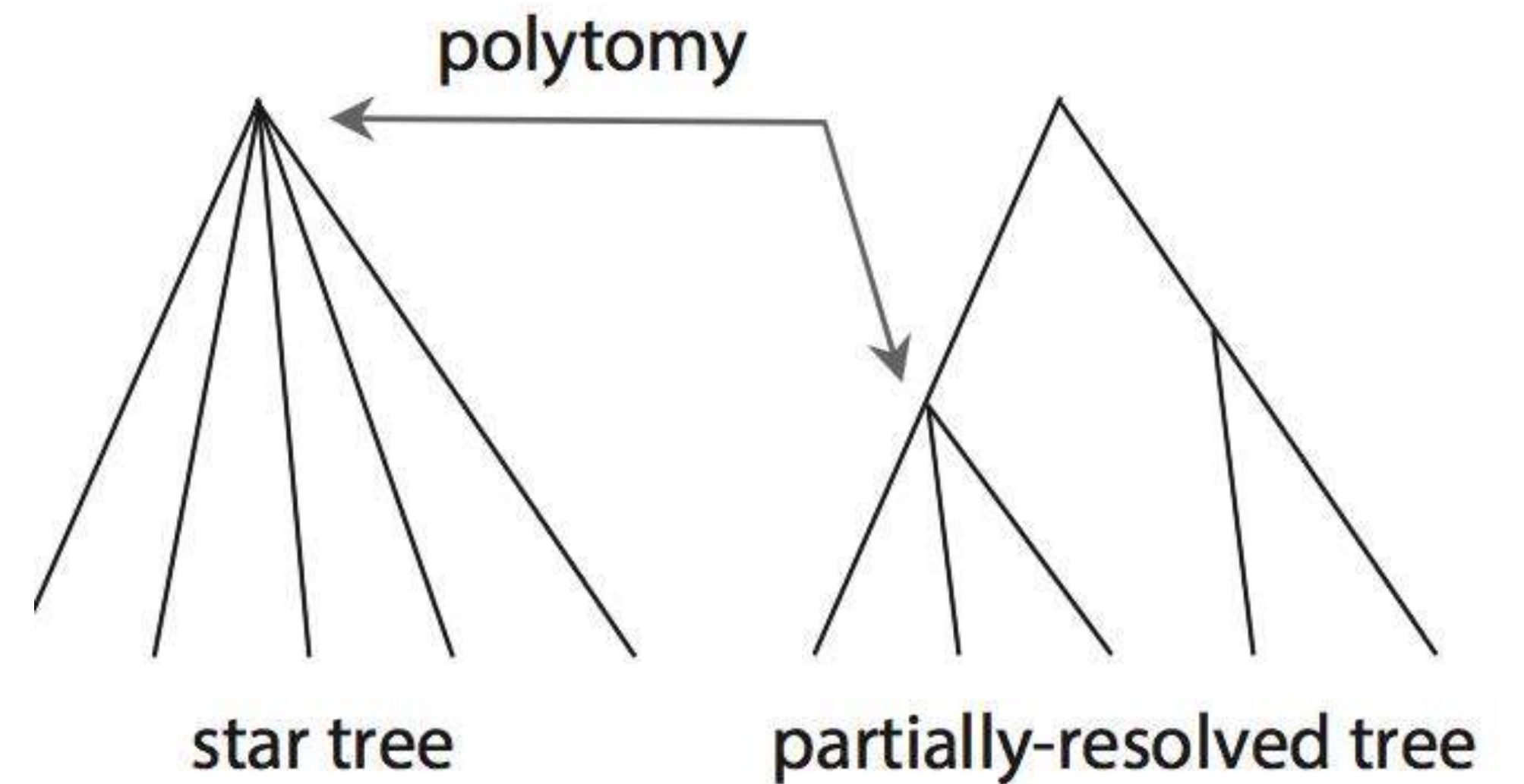


fully-resolved tree

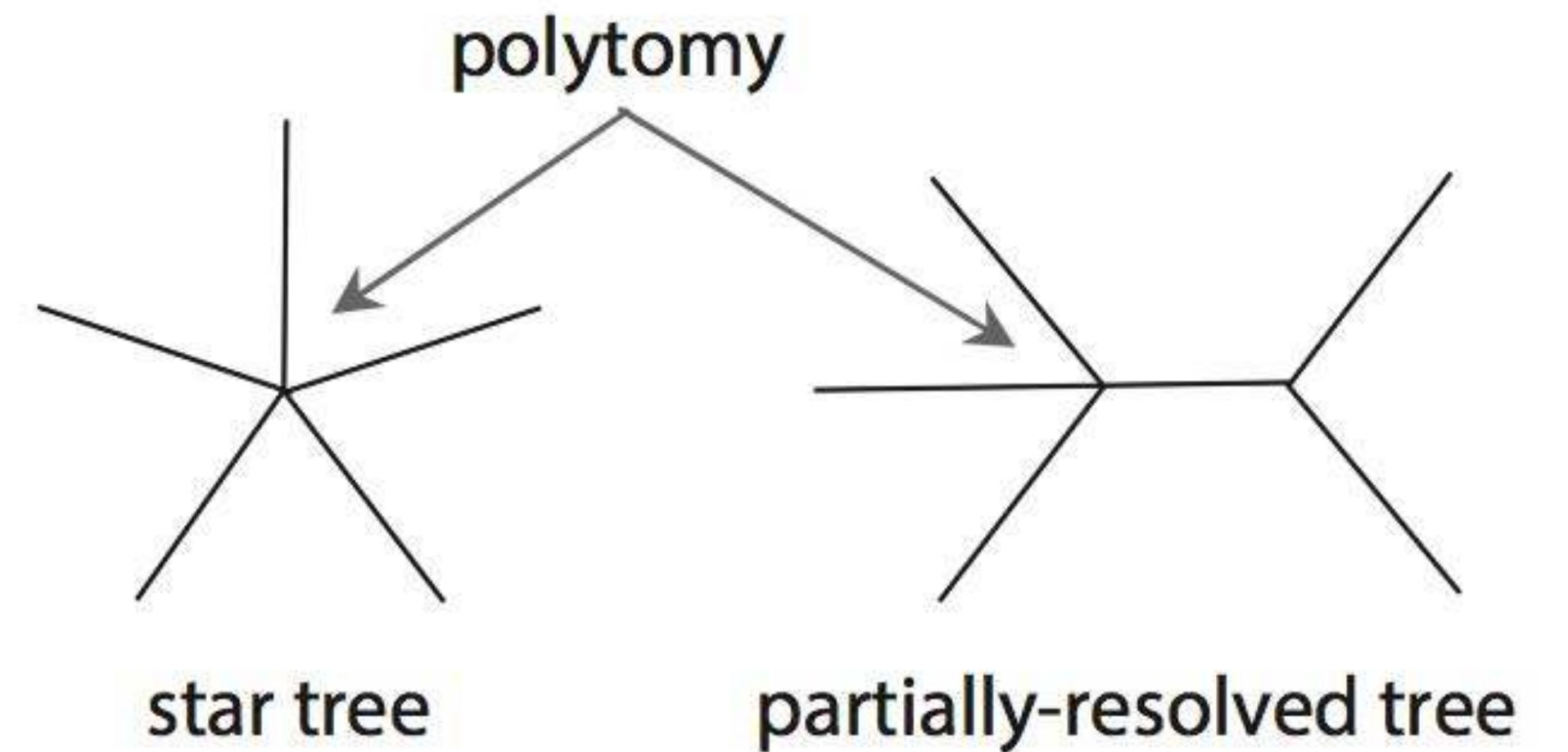


fully-resolved tree

Unresolved Phylogeny



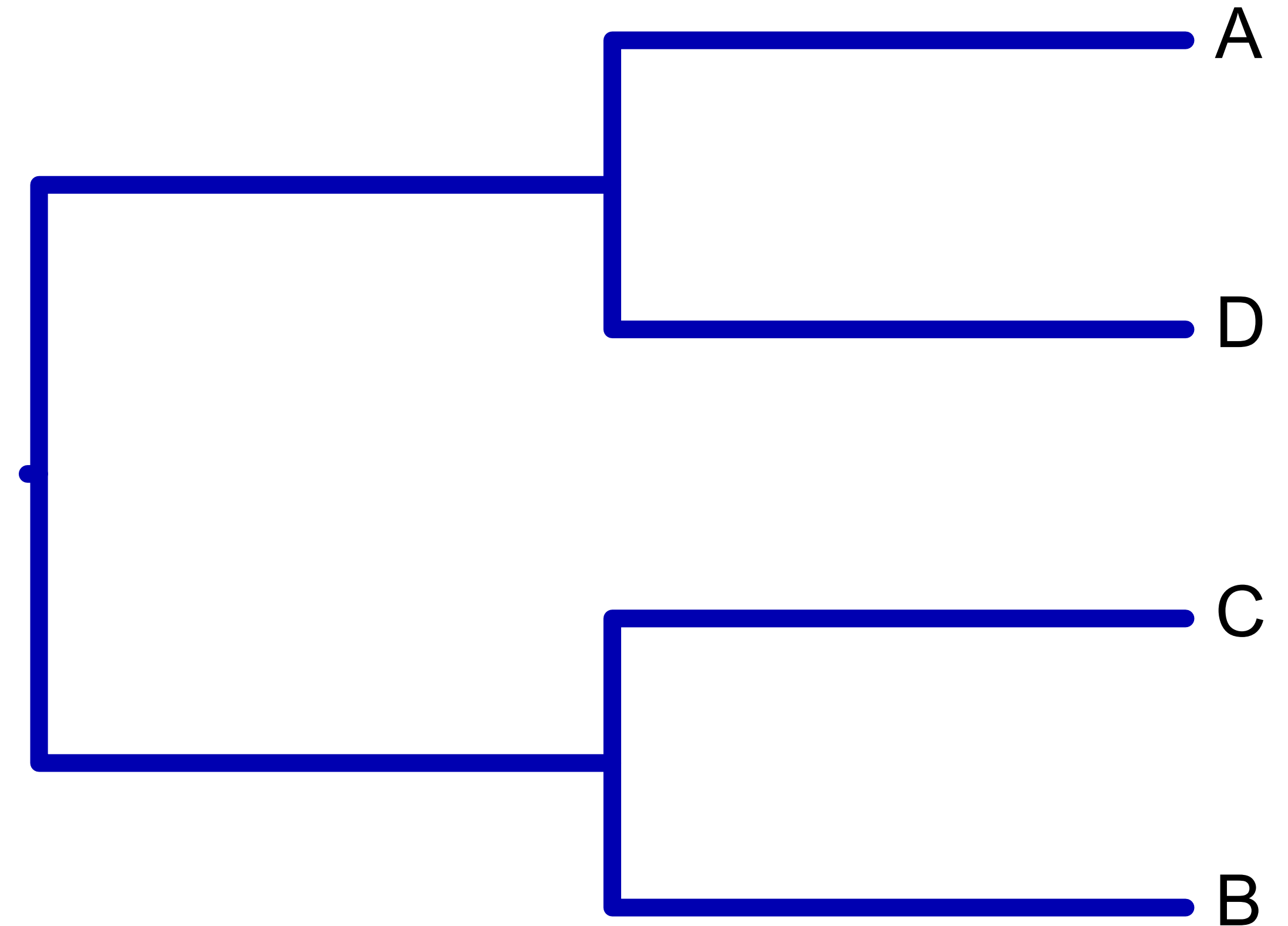
polytomies allow us to represent unresolved nodes



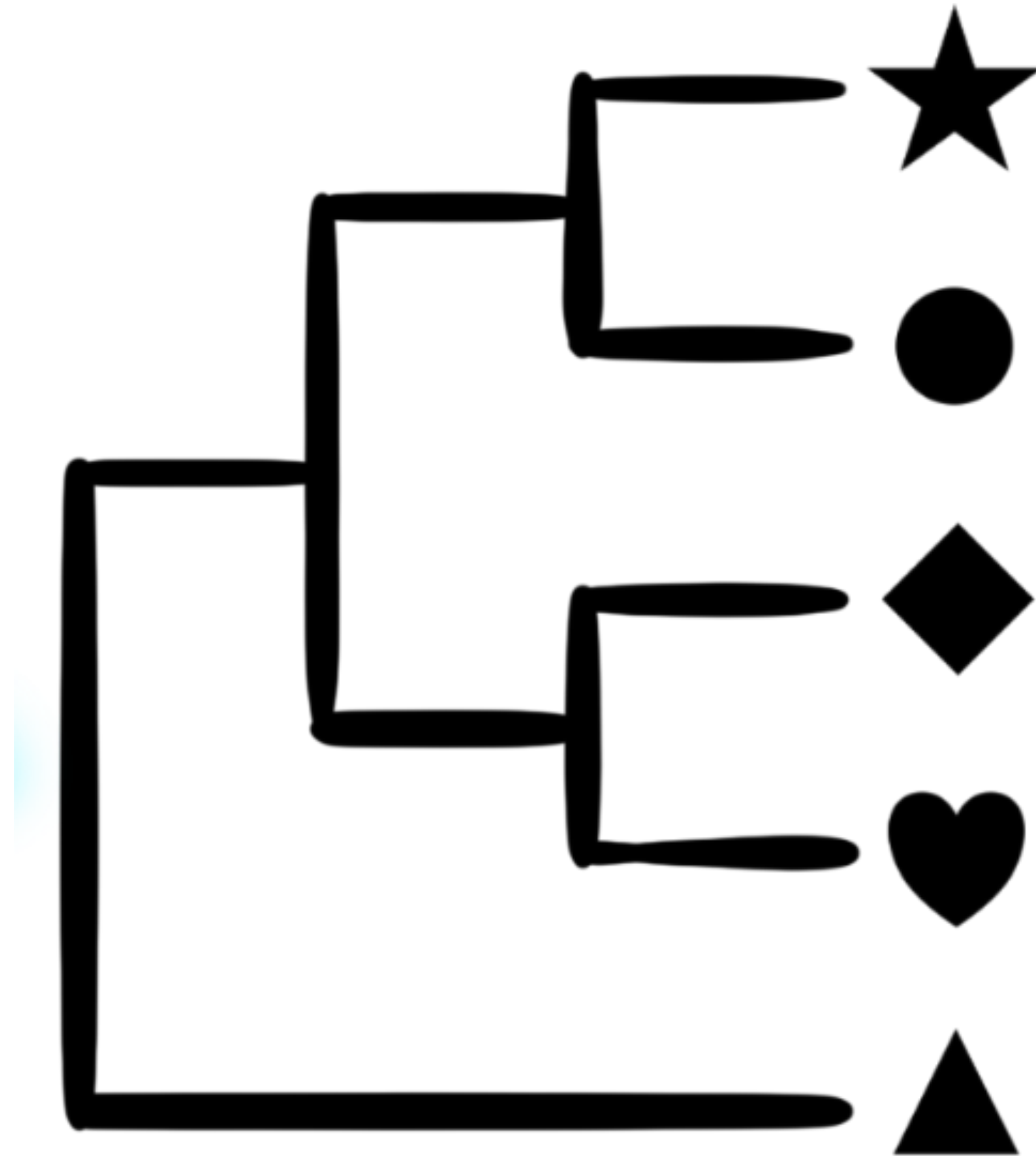
Representing Trees

$((A,D),(C,B));$

we can represent trees
using [Newick format](#), which
uses sets of nested
parentheses



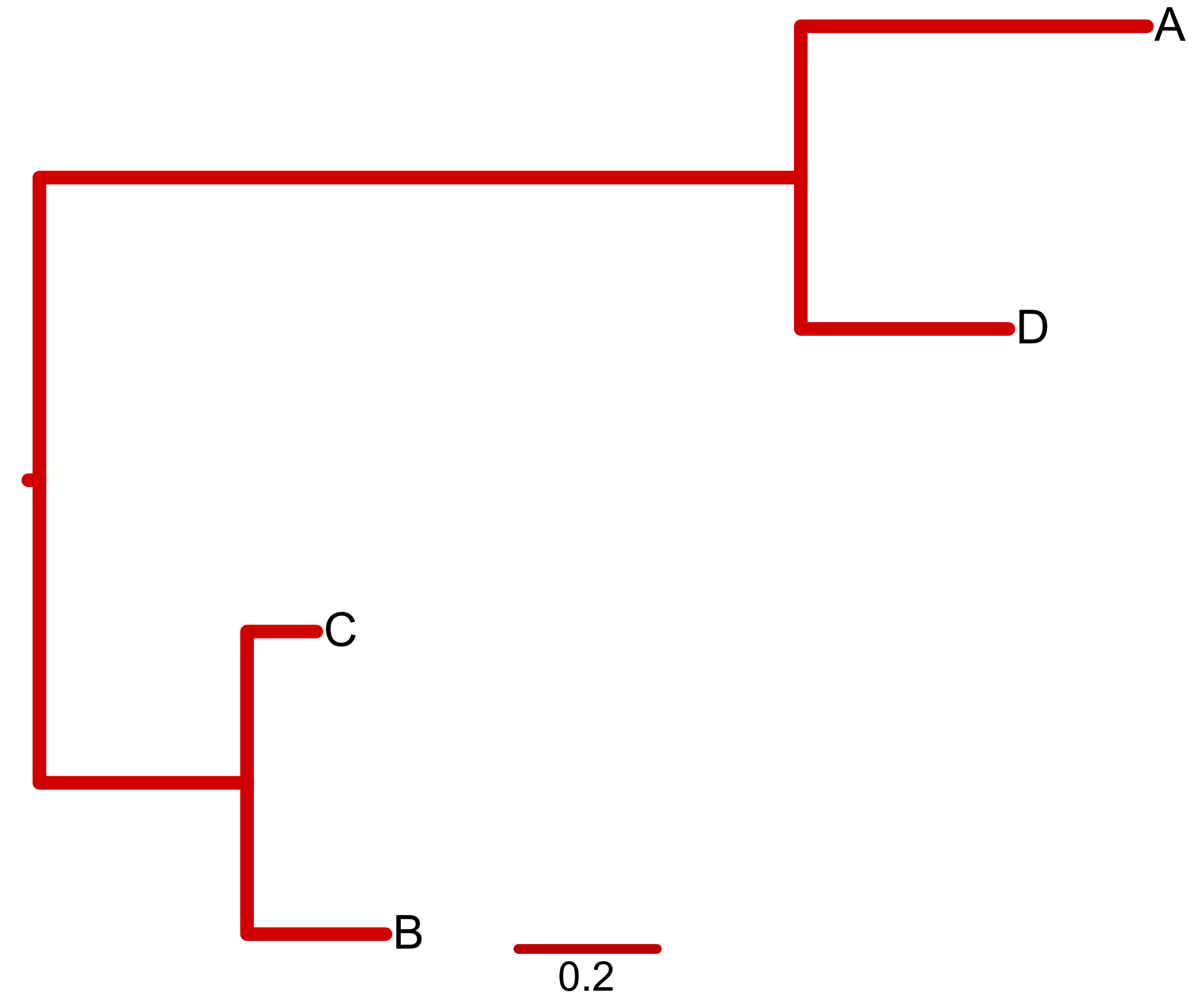
What is the Newick representation for this tree?



Representing Trees

this format also
accommodates branch
information

`((A:0.5,D:0.3):1.1,(C:0.1,B:0.2):0.3);`

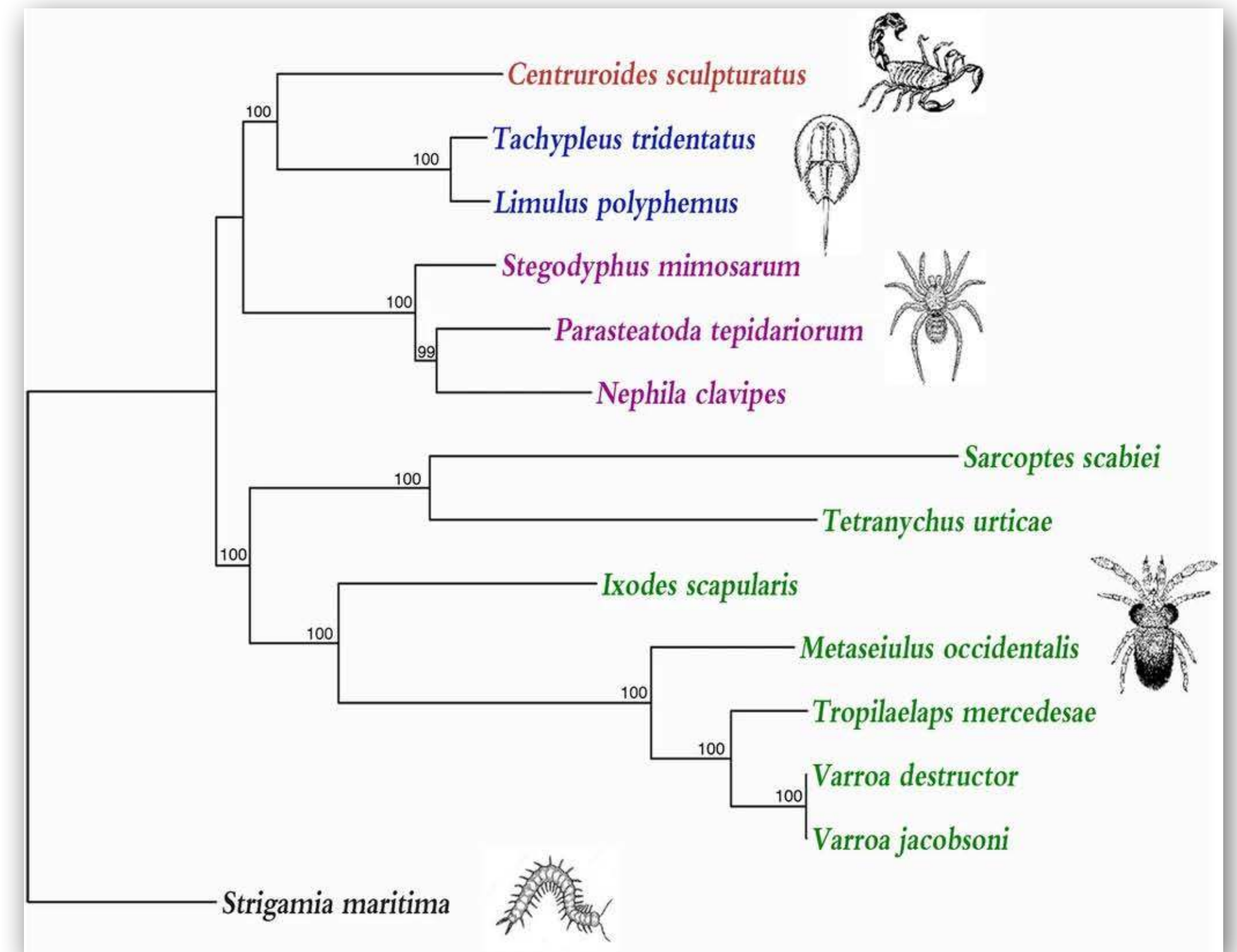
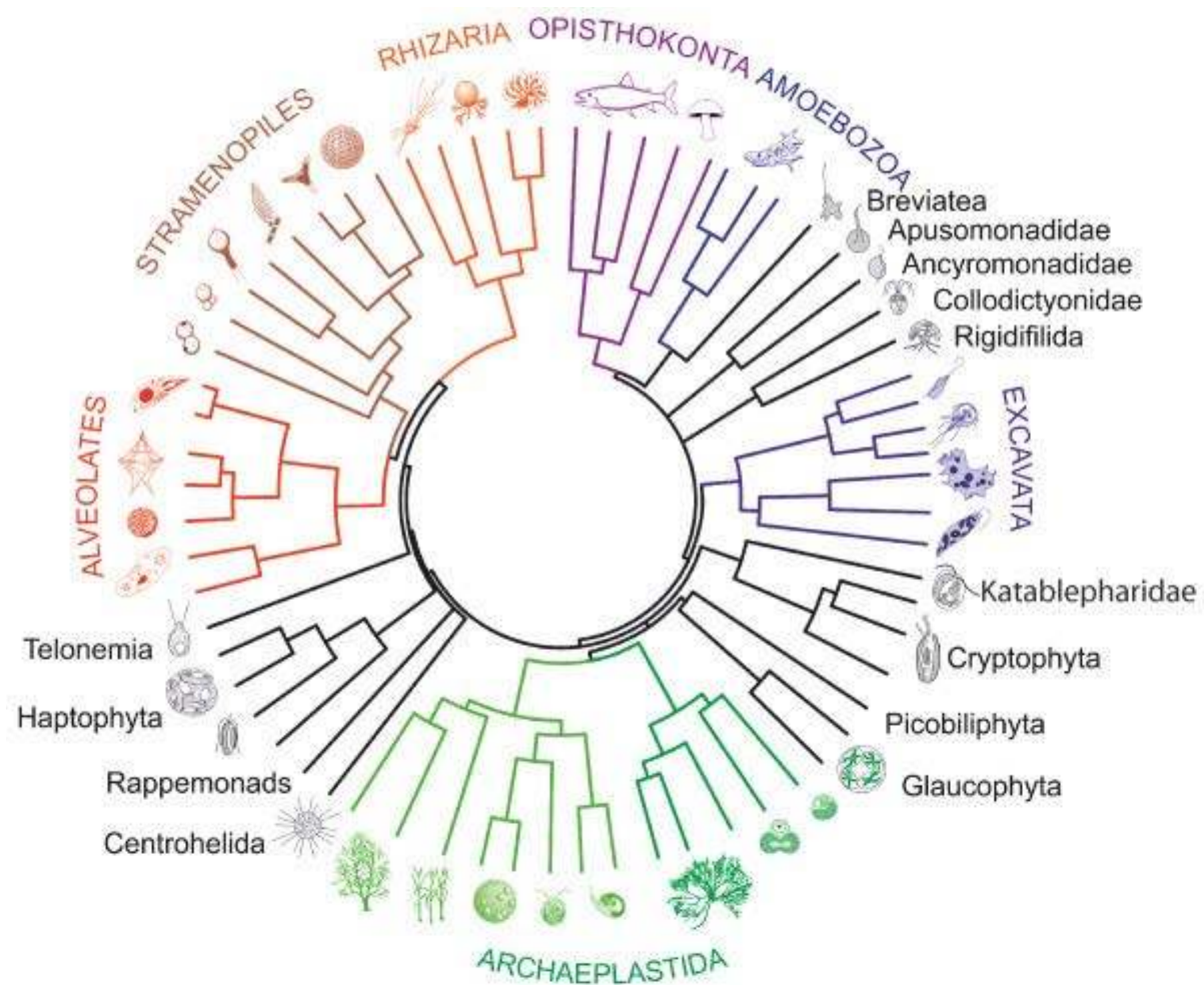


Draw the tree

$(((A, C), D), (B, E)), F);$

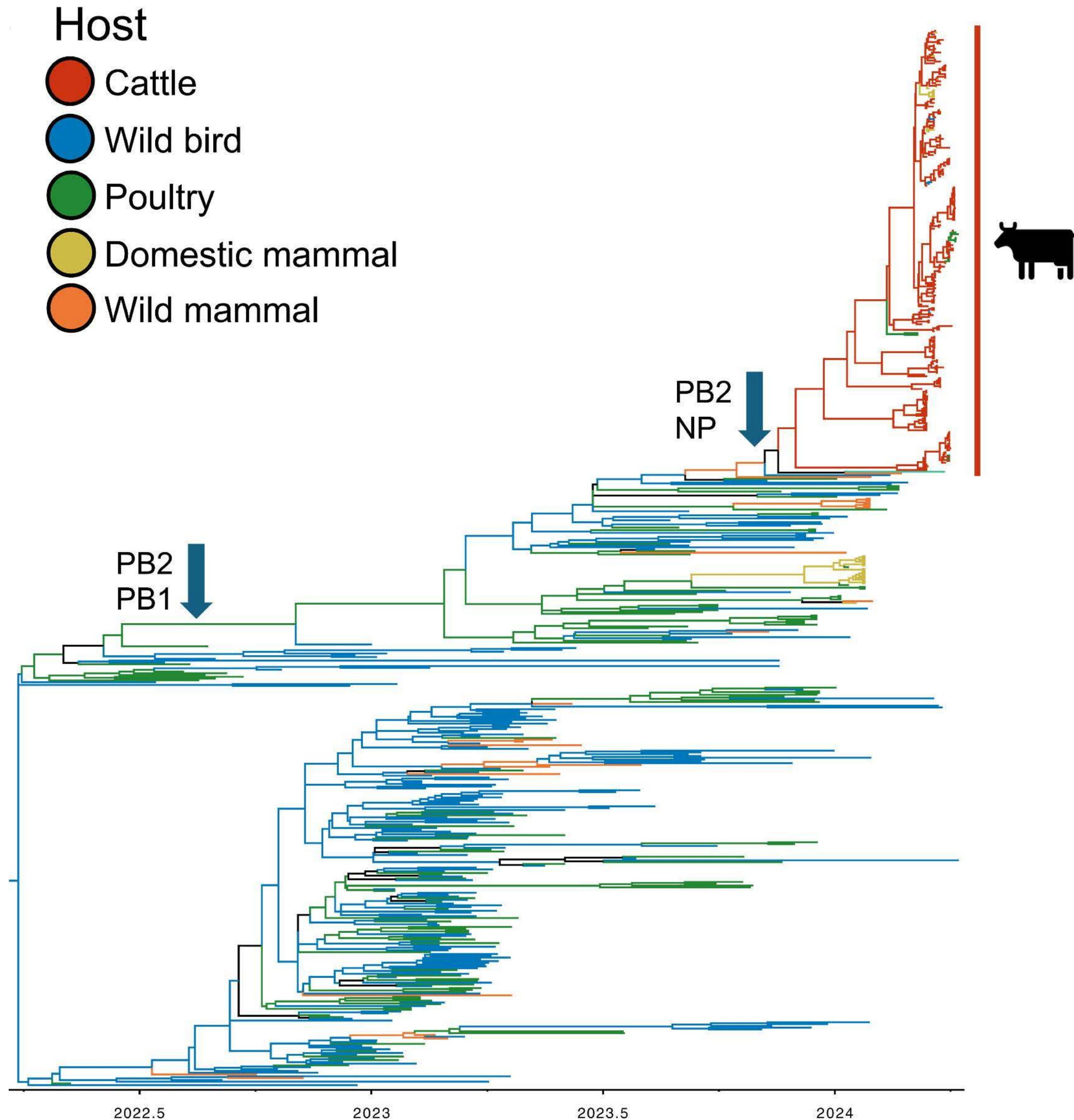
Phylogenetic Trees Depict Relatedness

Fungi are more closely related to Animals than they are to Plants



Horseshoe Crabs are Arachnids

Phylogenetic Trees Reveal Evolutionary Origins



Phylogenetic analysis of infectious diseases enable the study of the transmission dynamics

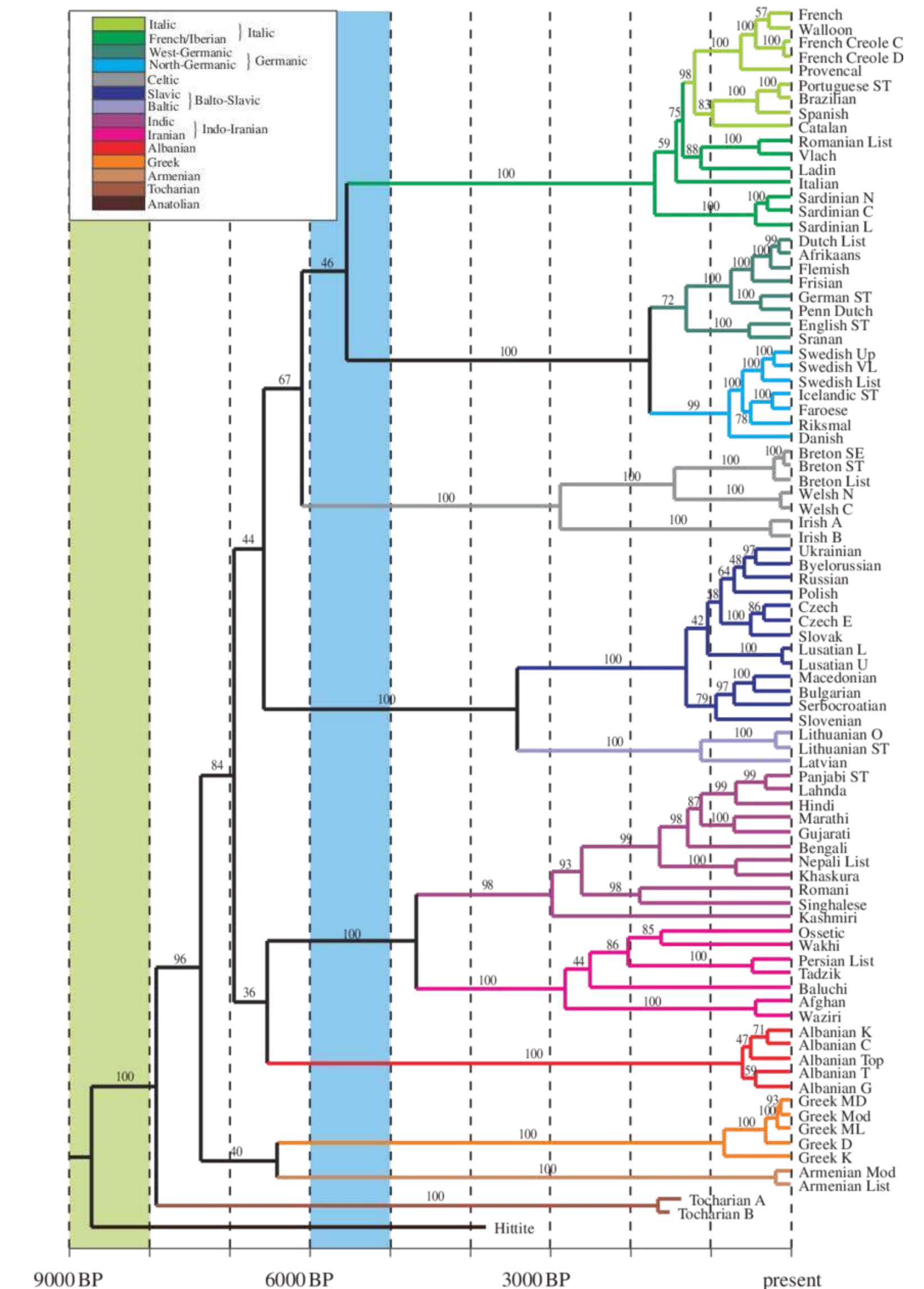
This is a phylogeny H5N1 sequences collected in North America Dec 2021 to April 2024

Highly pathogenic avian flu spread in dairy cattle after a single spillover event from wild birds

Phylogenetic Trees Expose Shared History

Phylogenetic methods are used to understand the history of human cultural evolution

This tree shows the diversification of Indo-European languages over time

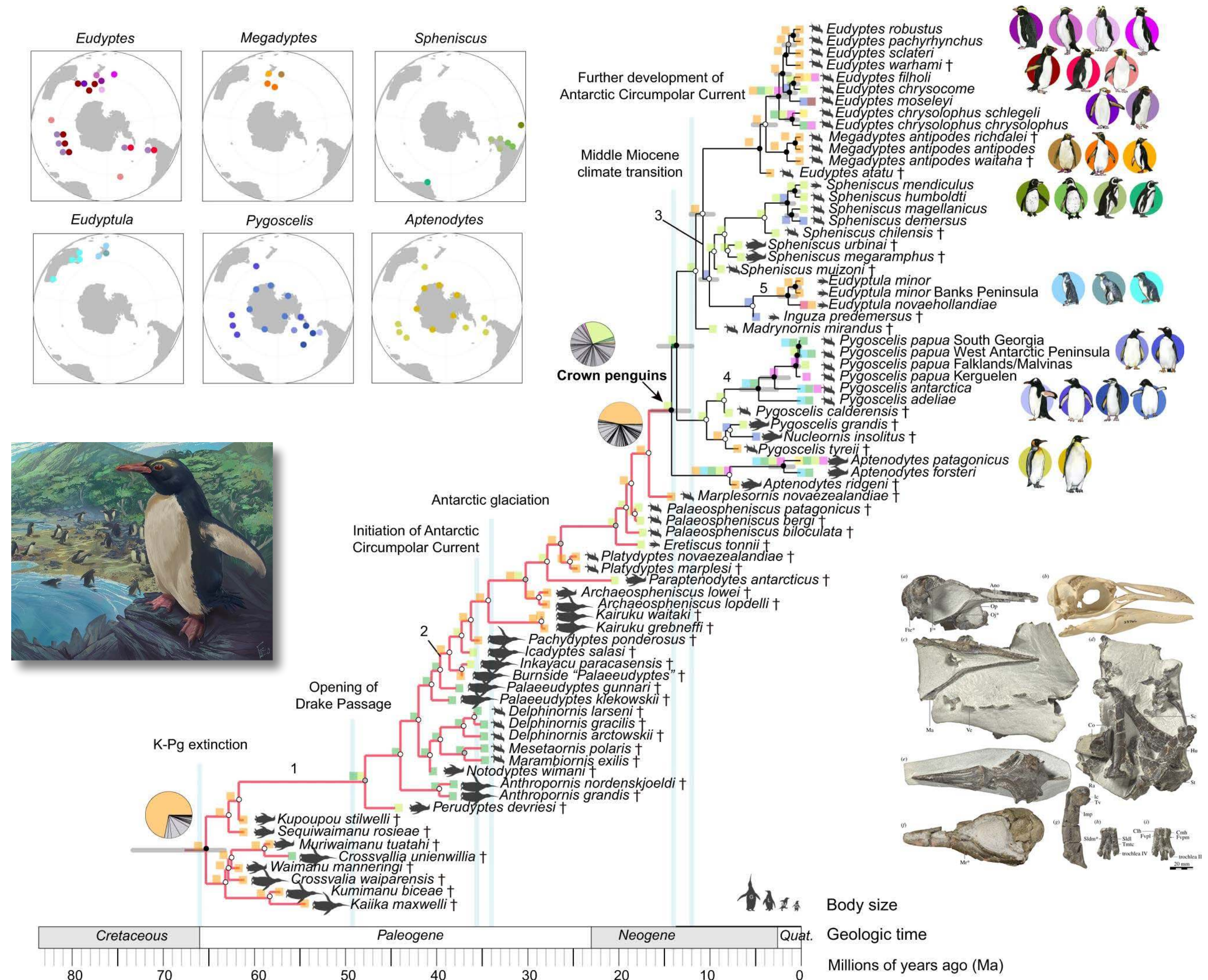


Phylogenetic Trees Illuminate Tempo & Mode

Phylogenetic methods provide statistical tests for macroevolution

Using several different data sources, we can learn about

- how genomes change over time
- shifts in speciation & extinction rates
- when traits evolved that enabled occupation of new ecological niches
- the geographical history of a group of species



Inferring Phylogenies

to estimate a phylogeny,
we need to start with
data

the data are observations
of character states for a
set of taxa

taxa	character 1	character 2	character 3
T1	pointed	blue	present
T2	pointed	blue	present
T3	round	blue	absent
T4	round	black	absent

a column in the matrix is a
character

the form that character takes
is its **state**

Inferring Phylogenies

to estimate a phylogeny,
we need to start with
data

the data are observations
of character states for a
set of taxa

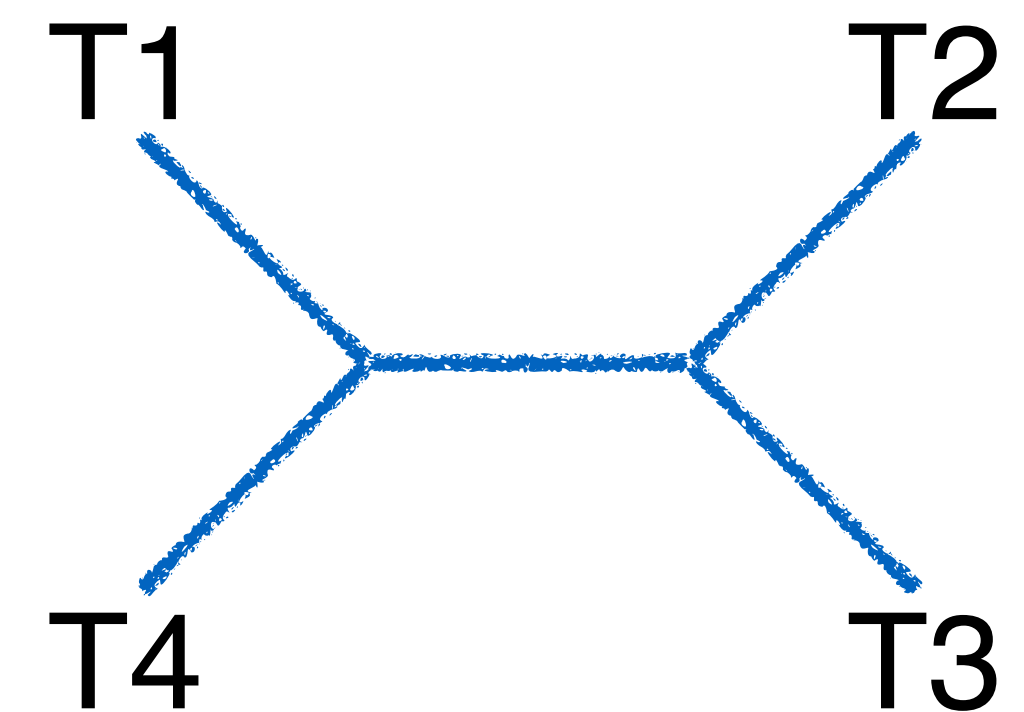
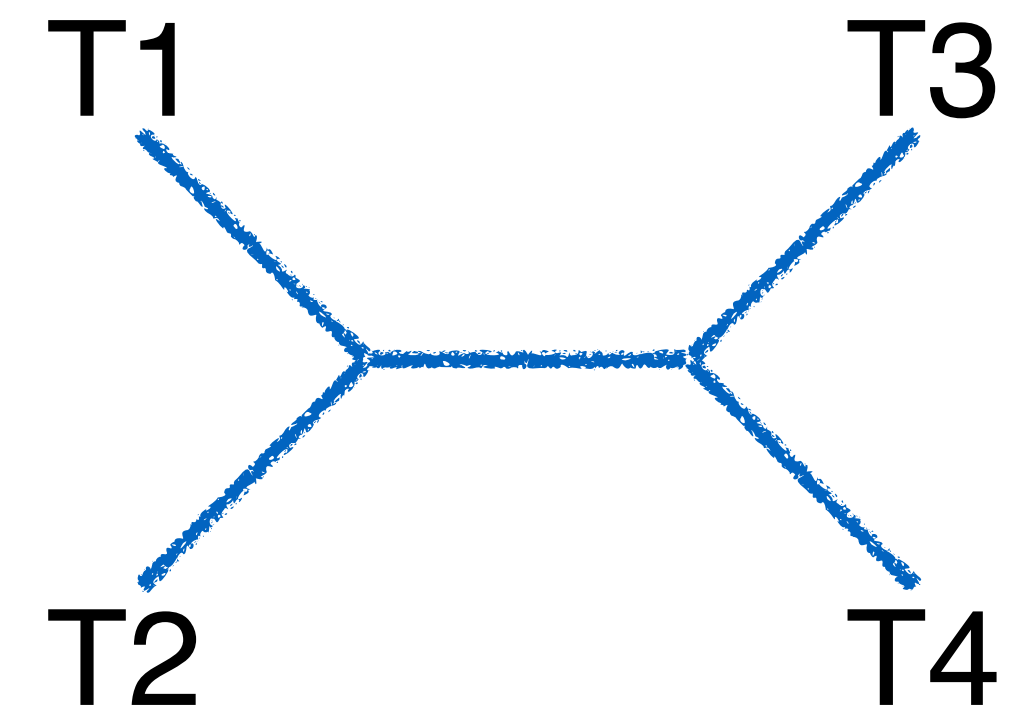
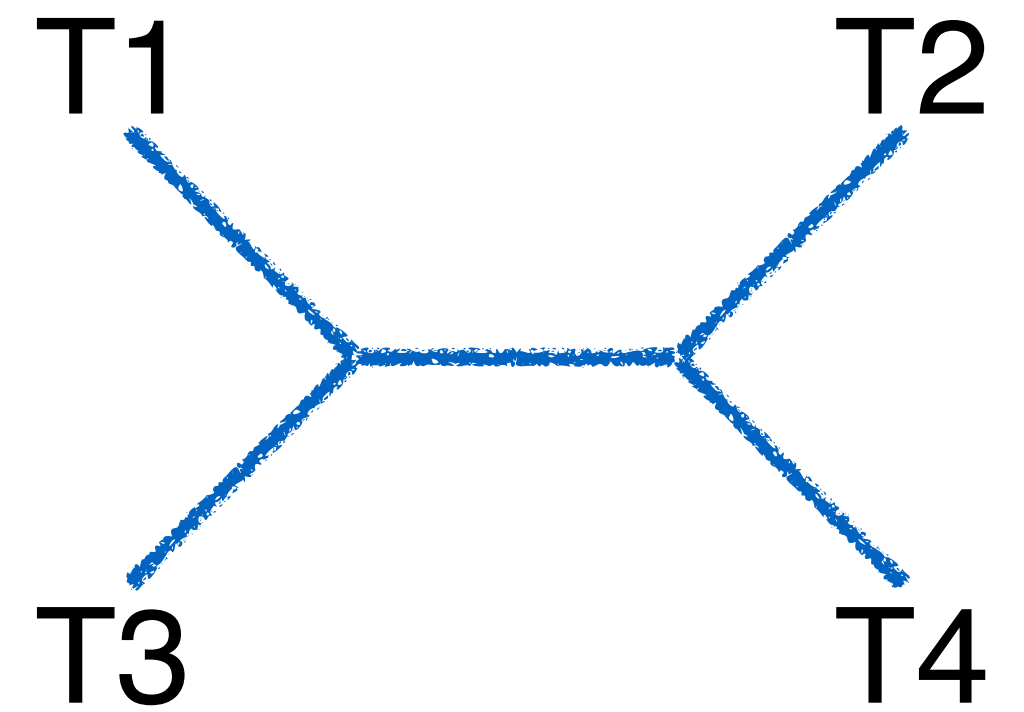
taxa	character 1	character 2	character 3
T1	A	T	C
T2	A	T	T
T3	G	T	G
T4	G	T	G

discrete characters can be
molecular or morphological

Inferring Phylogenies

we infer phylogenies by
evaluating tree topologies

for 4 taxa we can evaluate all
possible unrooted topologies
(there are only 3)



How many trees?

n	Unrooted trees (U_n)
3	1
4	3
5	15
6	105
7	945
8	10,395
9	135,135
10	2,027,025
20	$\sim 2.22 \times 10^{20}$
50	$\sim 2.84 \times 10^{74}$

$$U_n = \frac{(2n - 5)!}{2^{n-3}(n - 3)!}$$

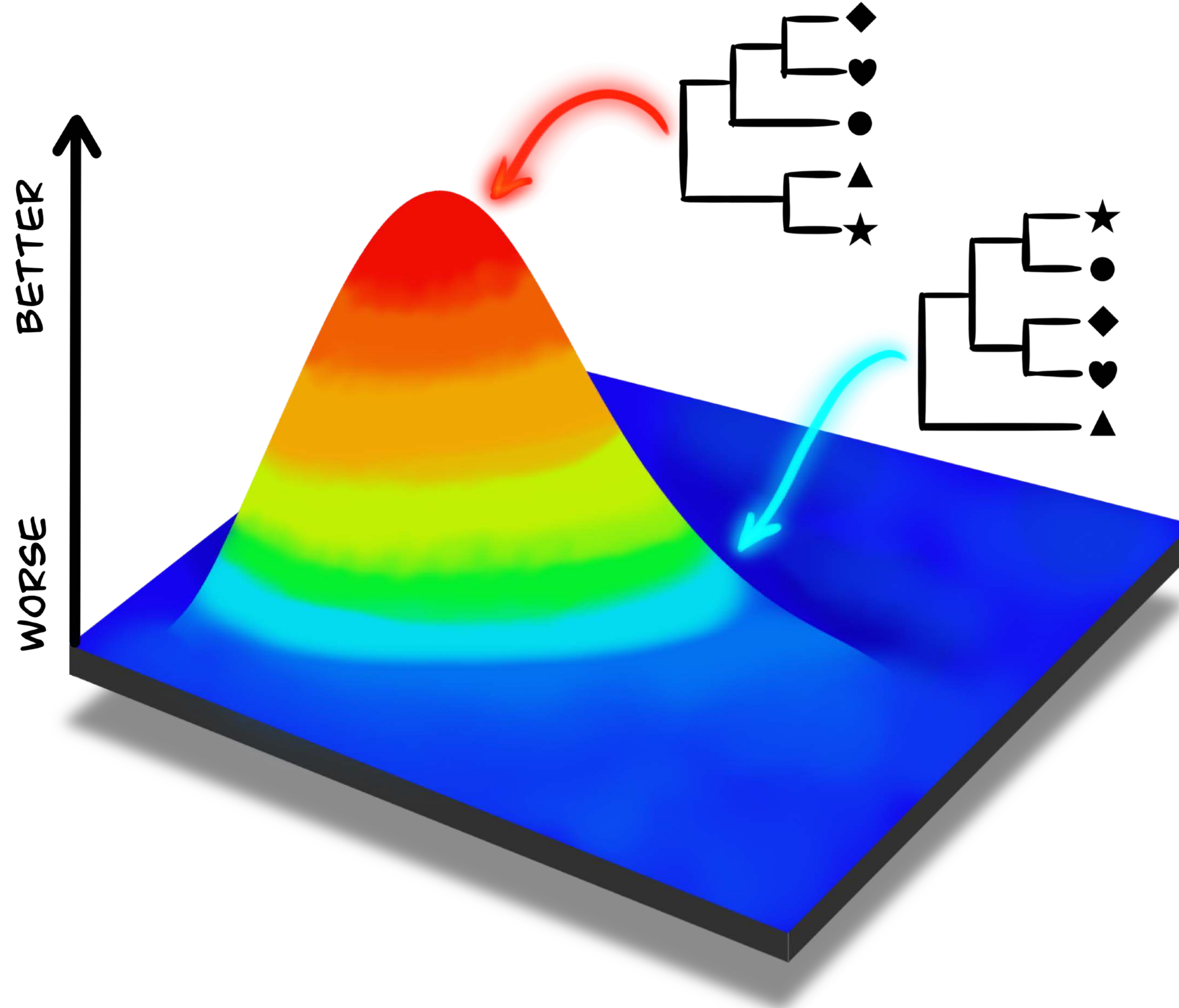
How many trees?

n	Unrooted trees (U_n)	Rooted trees (R_n)
3	1	3
4	3	15
5	15	105
6	105	945
7	945	10,395
8	10,395	135,135
9	135,135	2,027,025
10	2,027,025	34,459,425
20	$\sim 2.22 \times 10^{20}$	$\sim 8.20 \times 10^{21}$
50	$\sim 2.84 \times 10^{74}$	$\sim 2.75 \times 10^{76}$

$$R_n = \frac{(2n - 3)!}{2^{n-2}(n - 2)!}$$

at 51 taxa, the number of trees exceeds the number of particles in the observable universe

How to find the "best" tree?

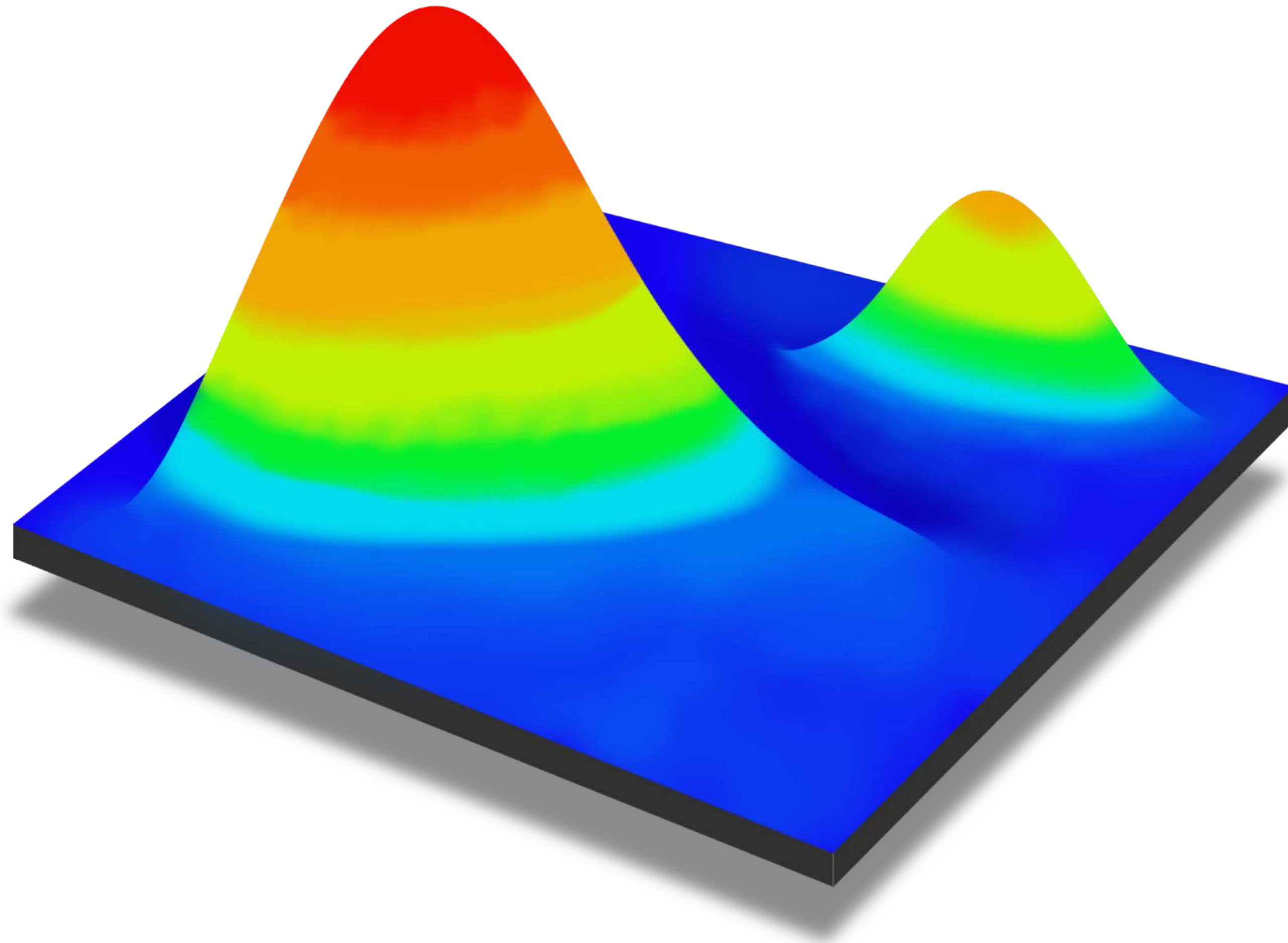


It depends on how you measure "best"

Table 3.2 Optimality criteria used for phylogeny reconstruction

Method	Criterion (tree score)
Maximum parsimony	Minimum number of changes, minimized over ancestral states
Maximum likelihood	Log likelihood score, optimized over branch lengths and model parameters
Minimum evolution	Tree length (sum of branch lengths, often estimated by least squares)
Bayesian	Posterior probability, calculated by integrating over branch lengths and substitution parameters

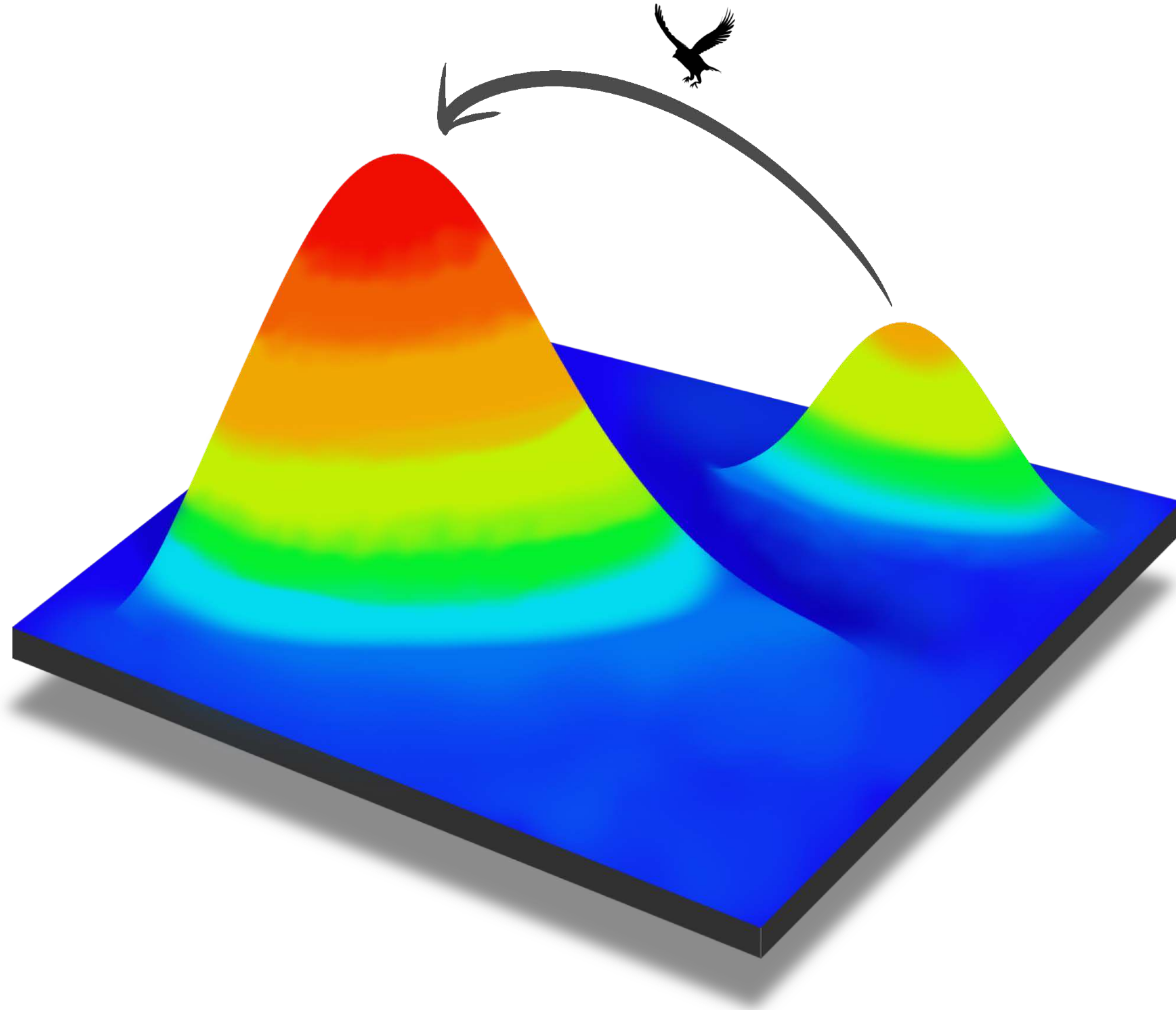
Searching Tree Topologies



a number of heuristic
tree-search algorithms
have been developed
([Yang 2014](#), Ch. 3)

these methods allow us
to evaluate just a
subset of the possible
trees

Searching Tree Topologies



importantly, we need tree-search methods that can find the global optimum

these approaches are useful for maximum parsimony, maximum likelihood, and Bayesian methods