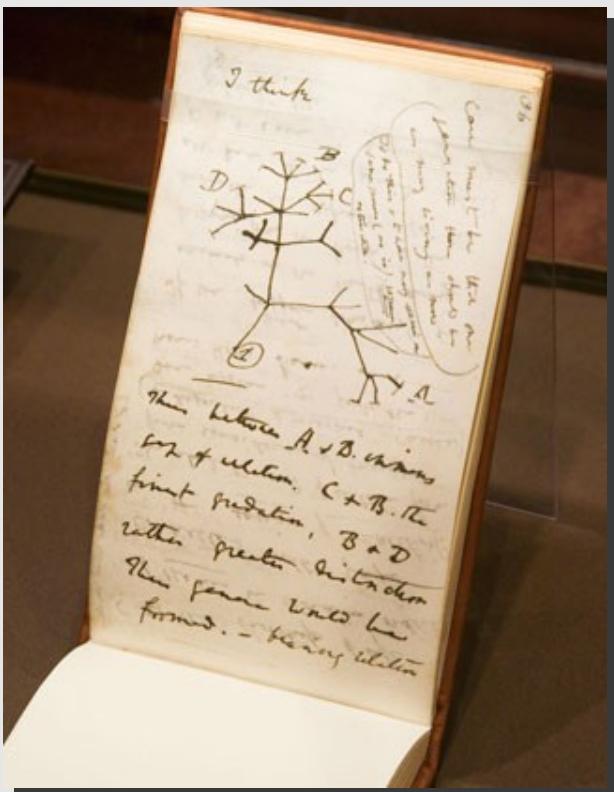


Phylogenetics



MY HOBBY:
EMBEDDING NP-COMPLETE PROBLEMS IN RESTAURANT ORDERS

CHOTCHKIES RESTAURANT

~ APPETIZERS ~

MIXED FRUIT	2.15
FRENCH FRIES	2.75
SIDE SALAD	3.35
HOT WINGS	3.55
MOZZARELLA STICKS	4.20
SAMPLER PLATE	5.80

~ SANDWICHES ~

BARBECUE	6.55
----------	------

WE'D LIKE EXACTLY \$15.05
WORTH OF APPETIZERS, PLEASE.

... EXACTLY? UHH ...

HERE, THESE PAPERS ON THE KNAPSACK
PROBLEM MIGHT HELP YOU OUT.

LISTEN, I HAVE SIX OTHER
TABLES TO GET TO -

- AS FAST AS POSSIBLE, OF COURSE. WANT
SOMETHING ON TRAVELING SALESMAN?

<http://xkcd.com/287/>

Brian O'Meara
EEB464 Fall 2018

Learning objectives

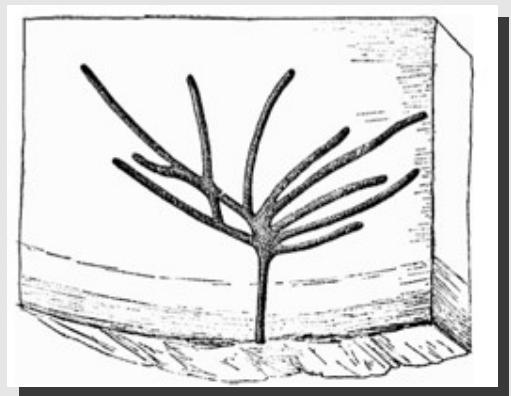
- Why bother with phylogenies
- Questions we can answer using these techniques



© Doug Stone

0.7 mm

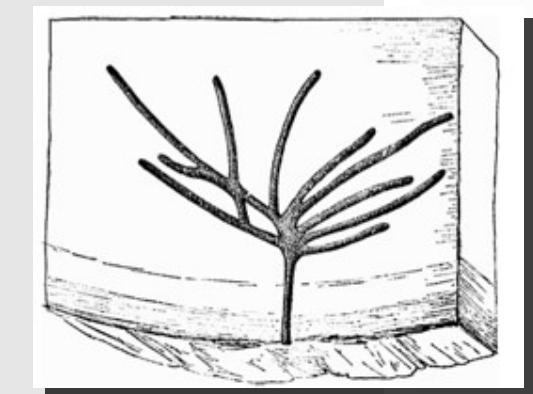
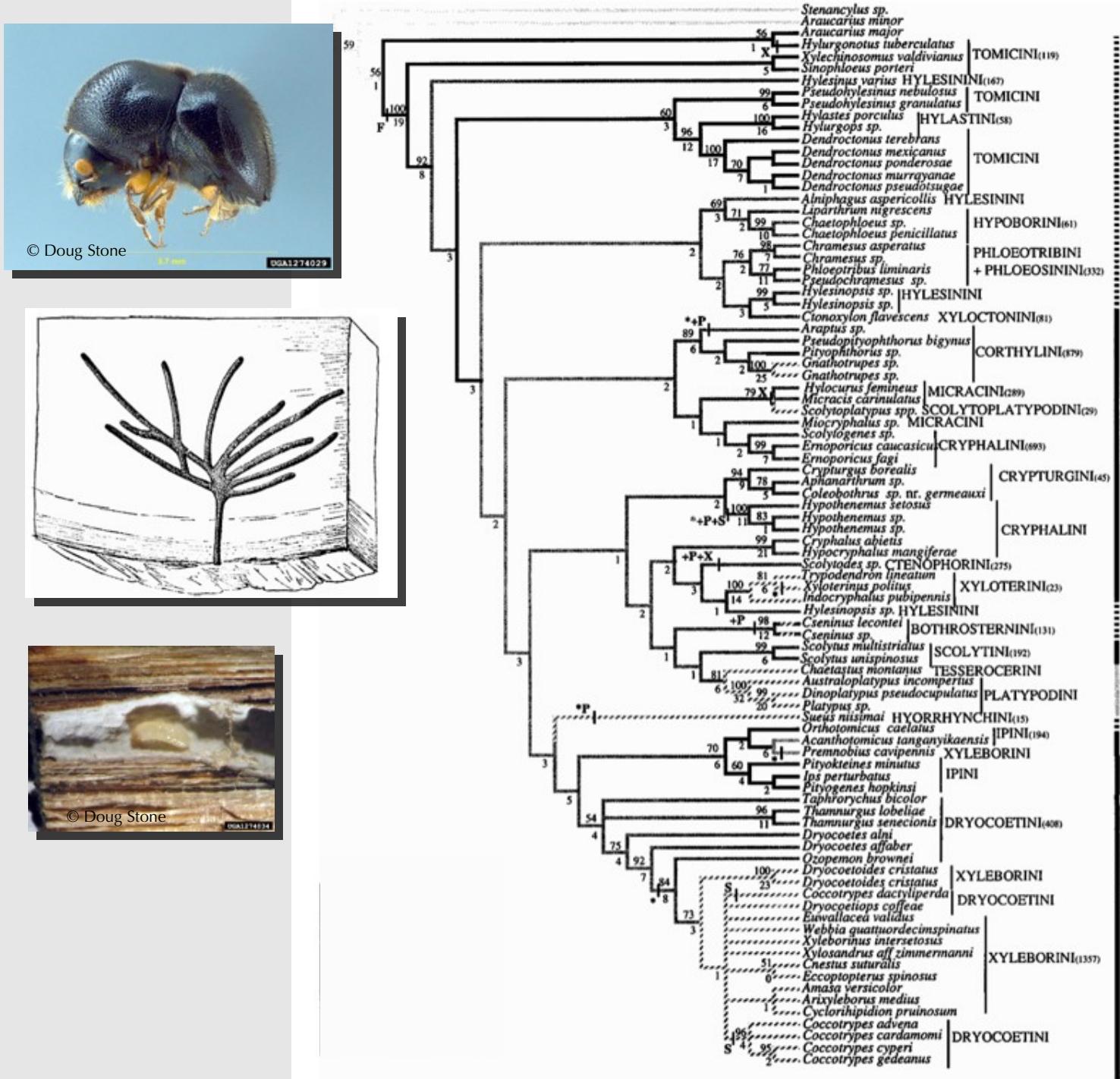
DGA1274029

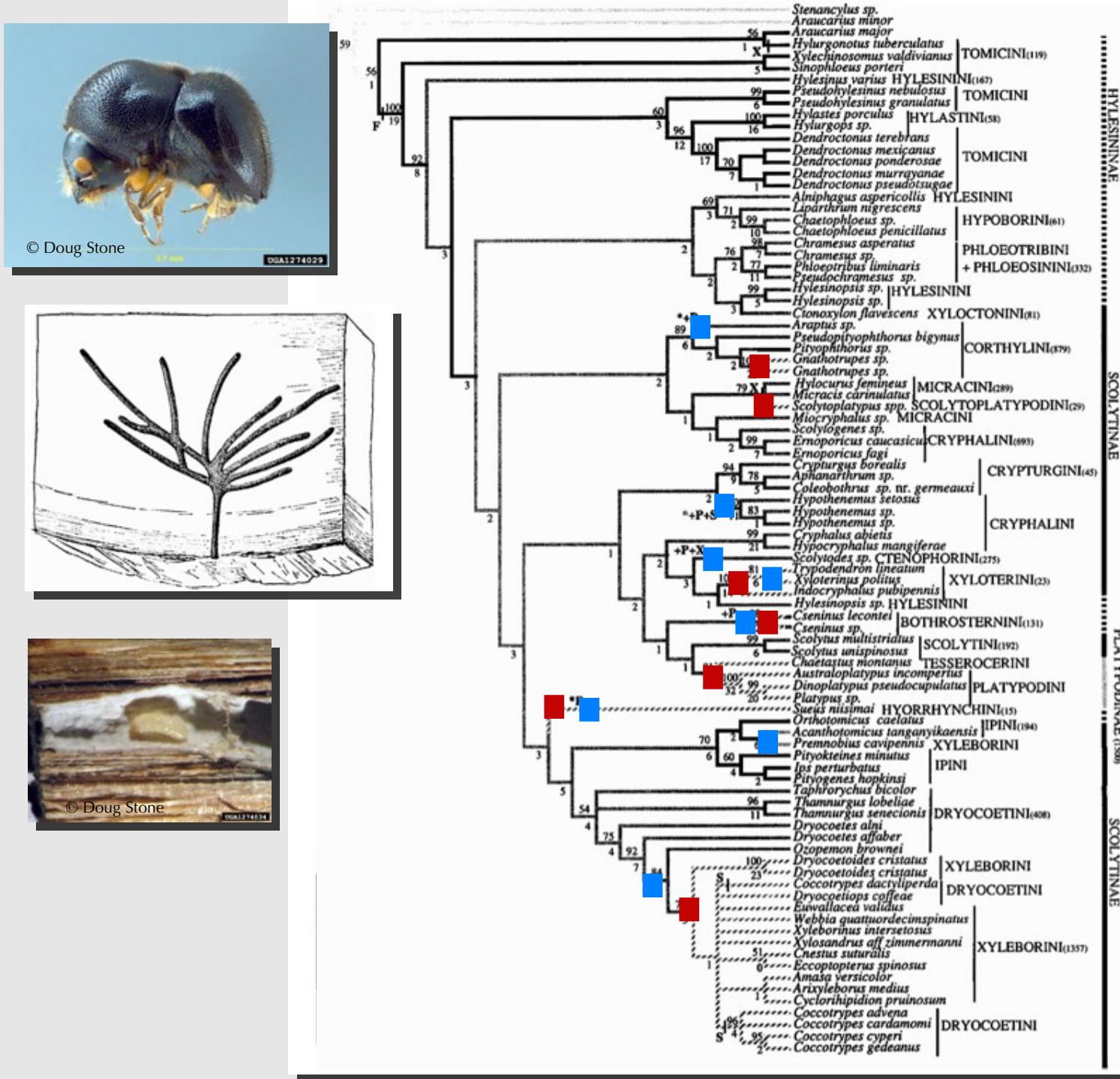


© Doug Stone

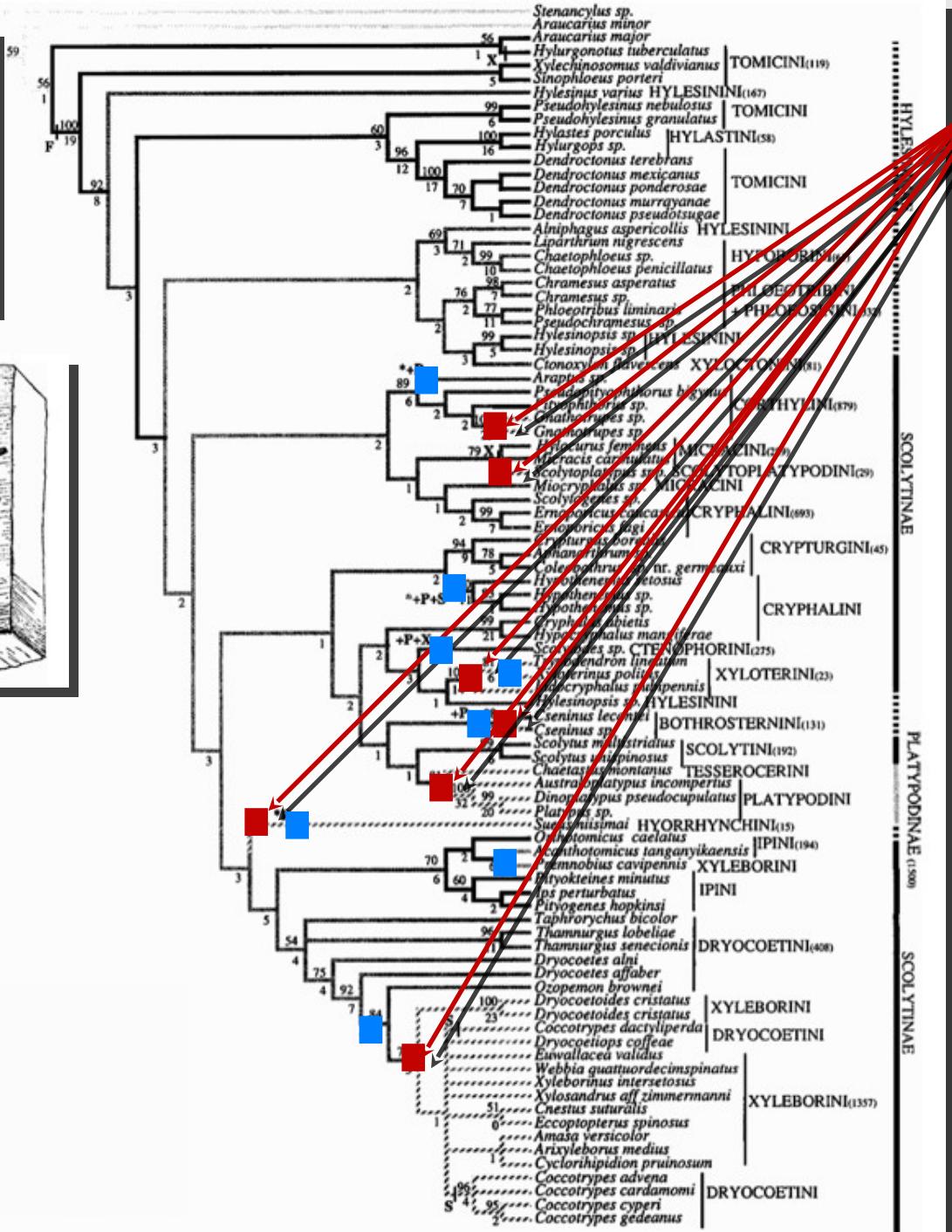
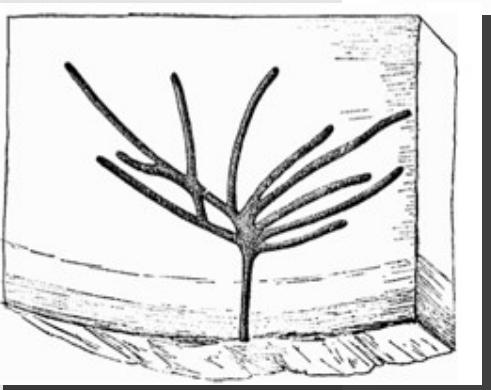
0.7 mm

DGA1274034

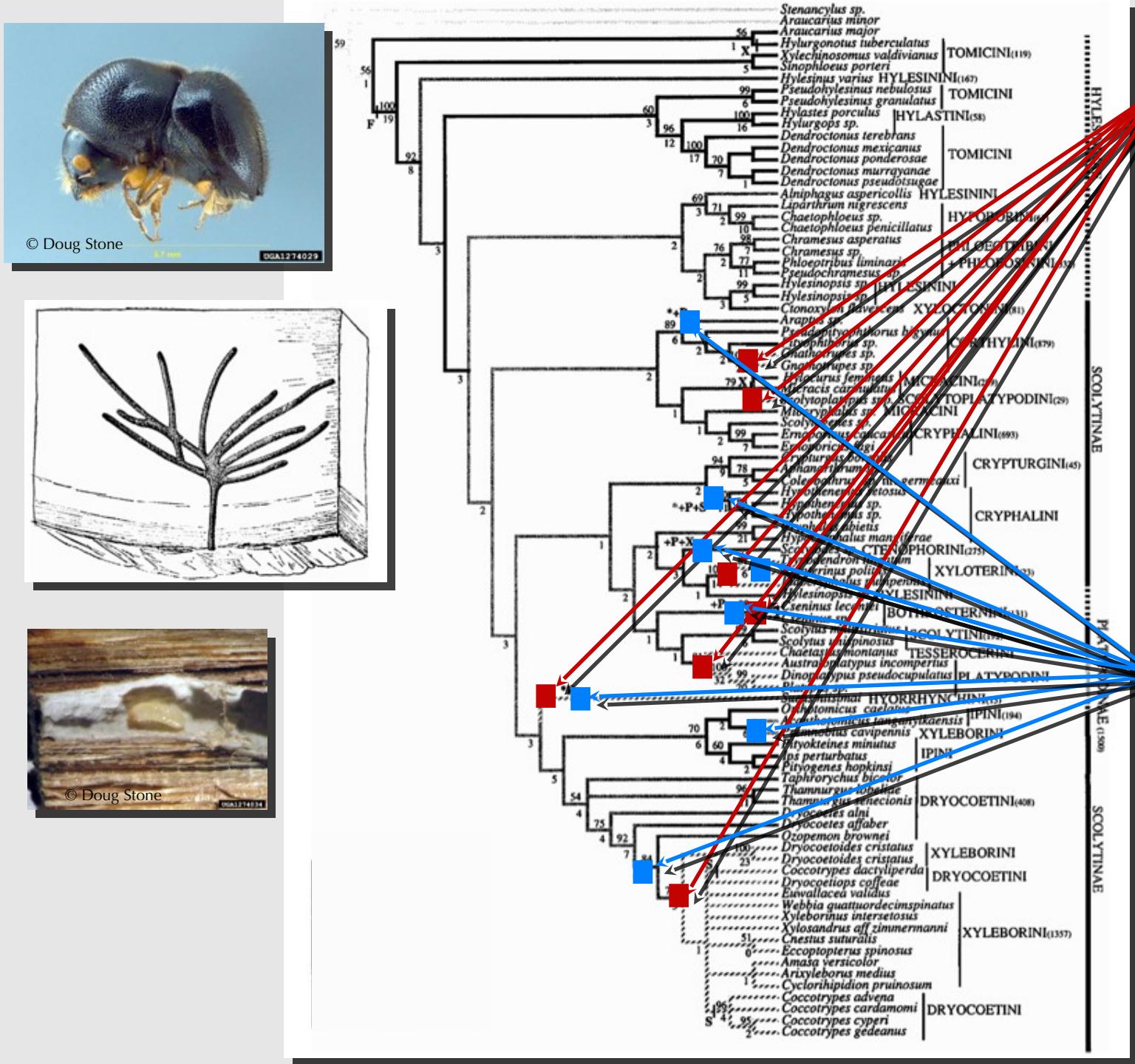




7 origins of agriculture



7 origins of agriculture



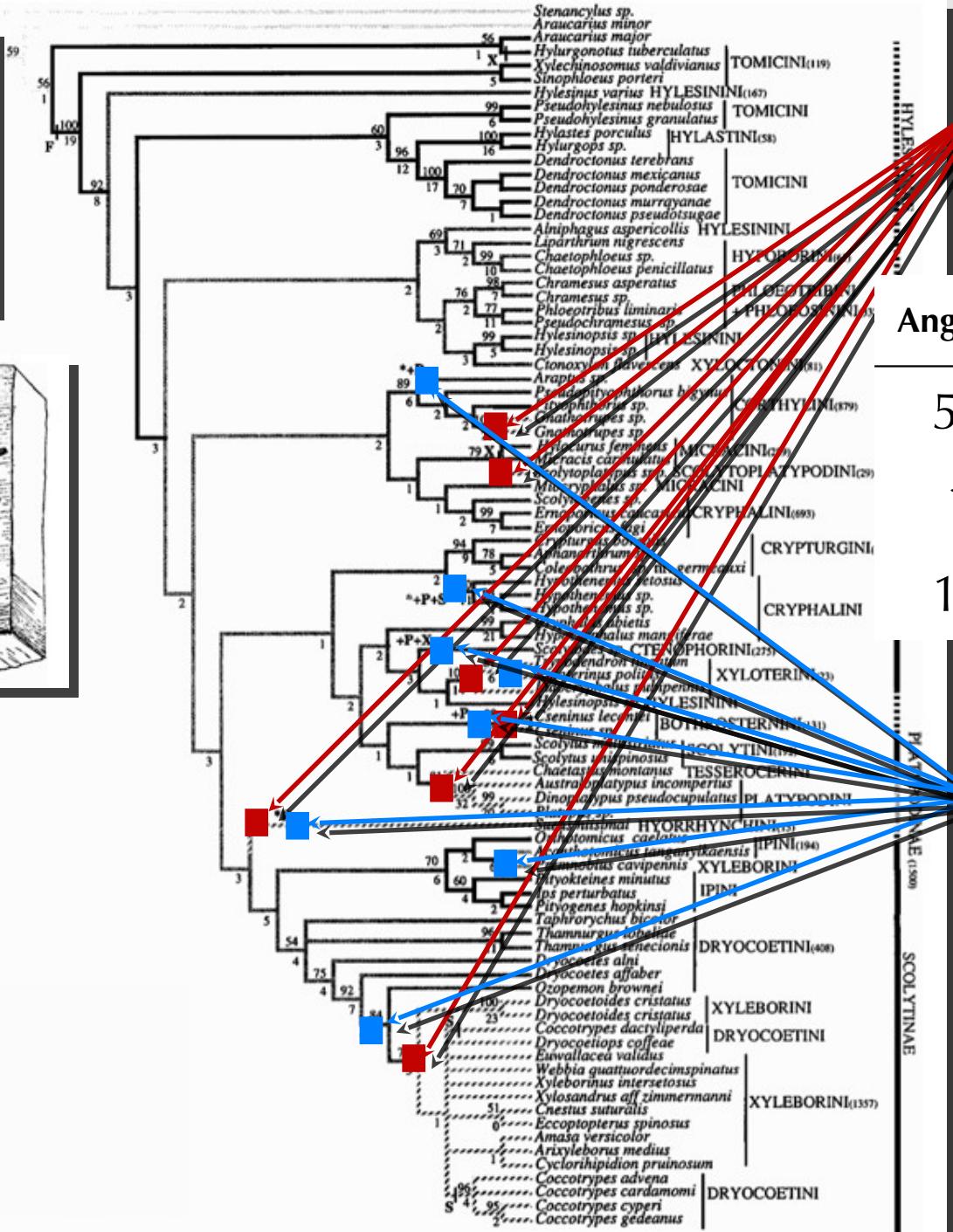
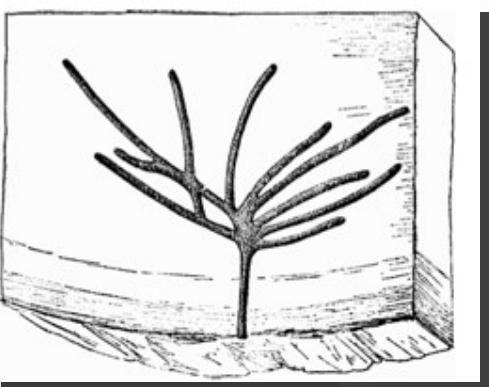
8 origins of inbreeding

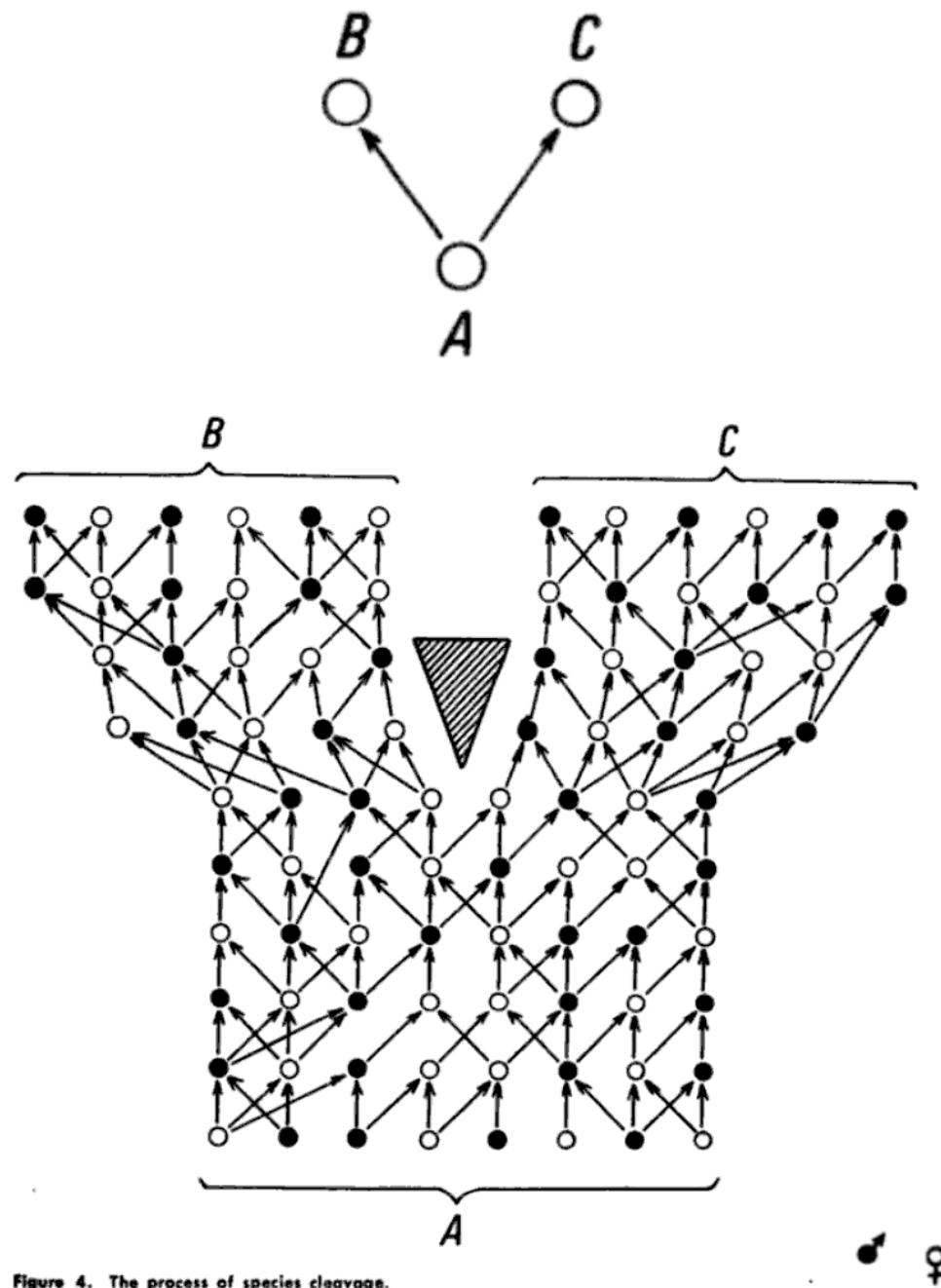
7 origins of agriculture

Angiosperm Conifer

5200	180
119	11
1500	195

8 origins of inbreeding





Willi Hennig,
Phylogenetic
Systematics, 1966.

Figure 4. The process of species cleavage.

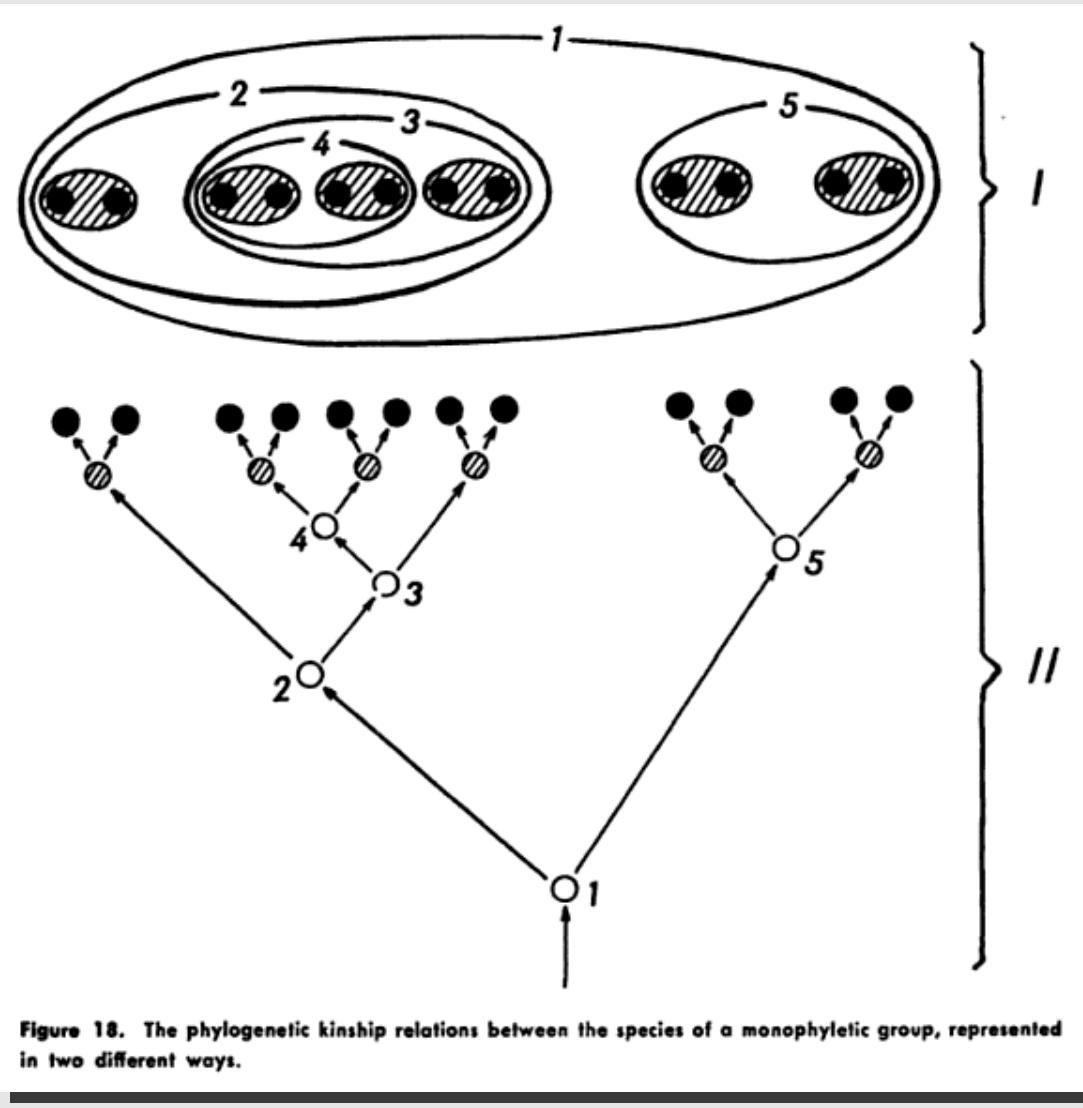
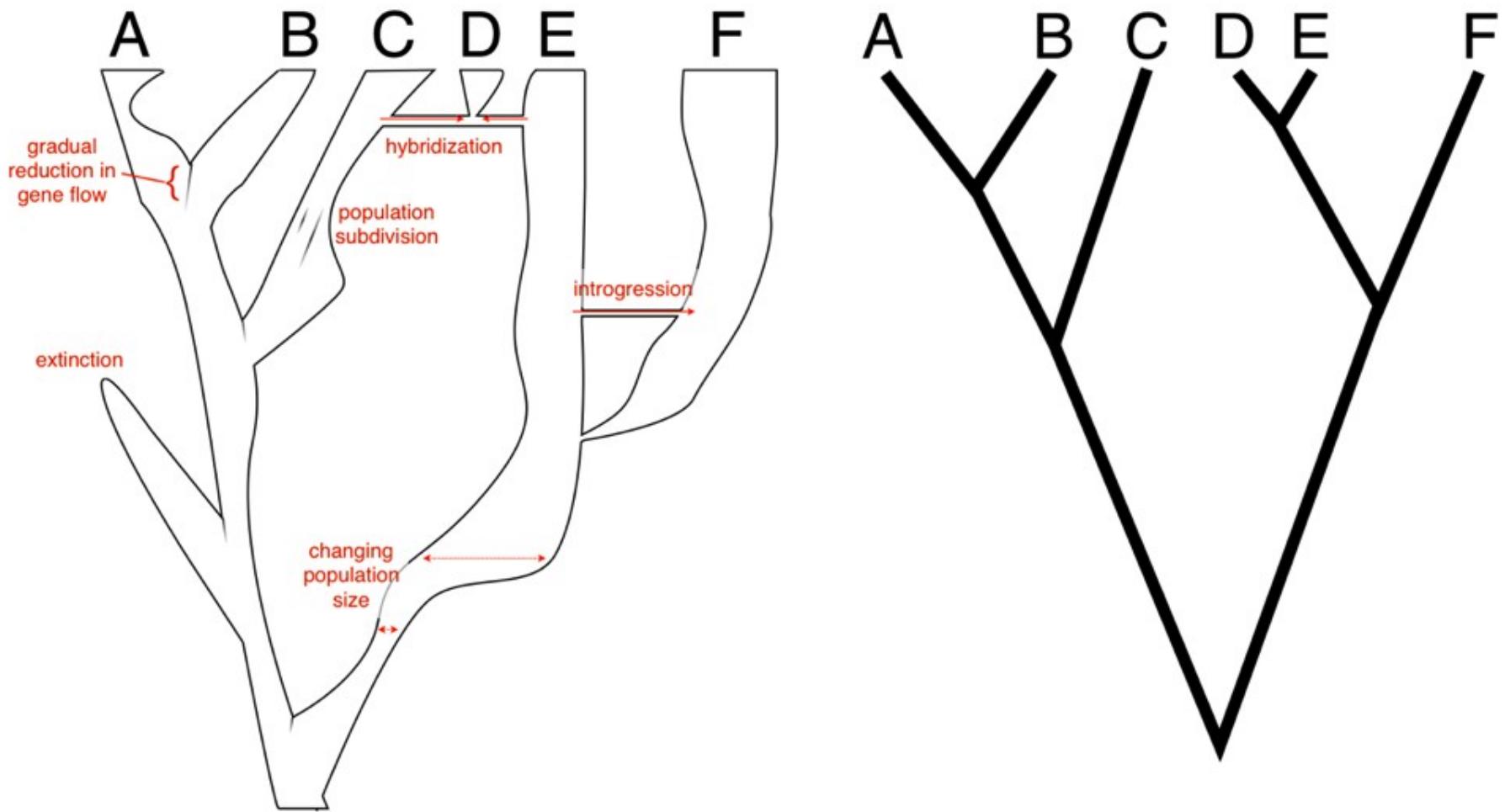
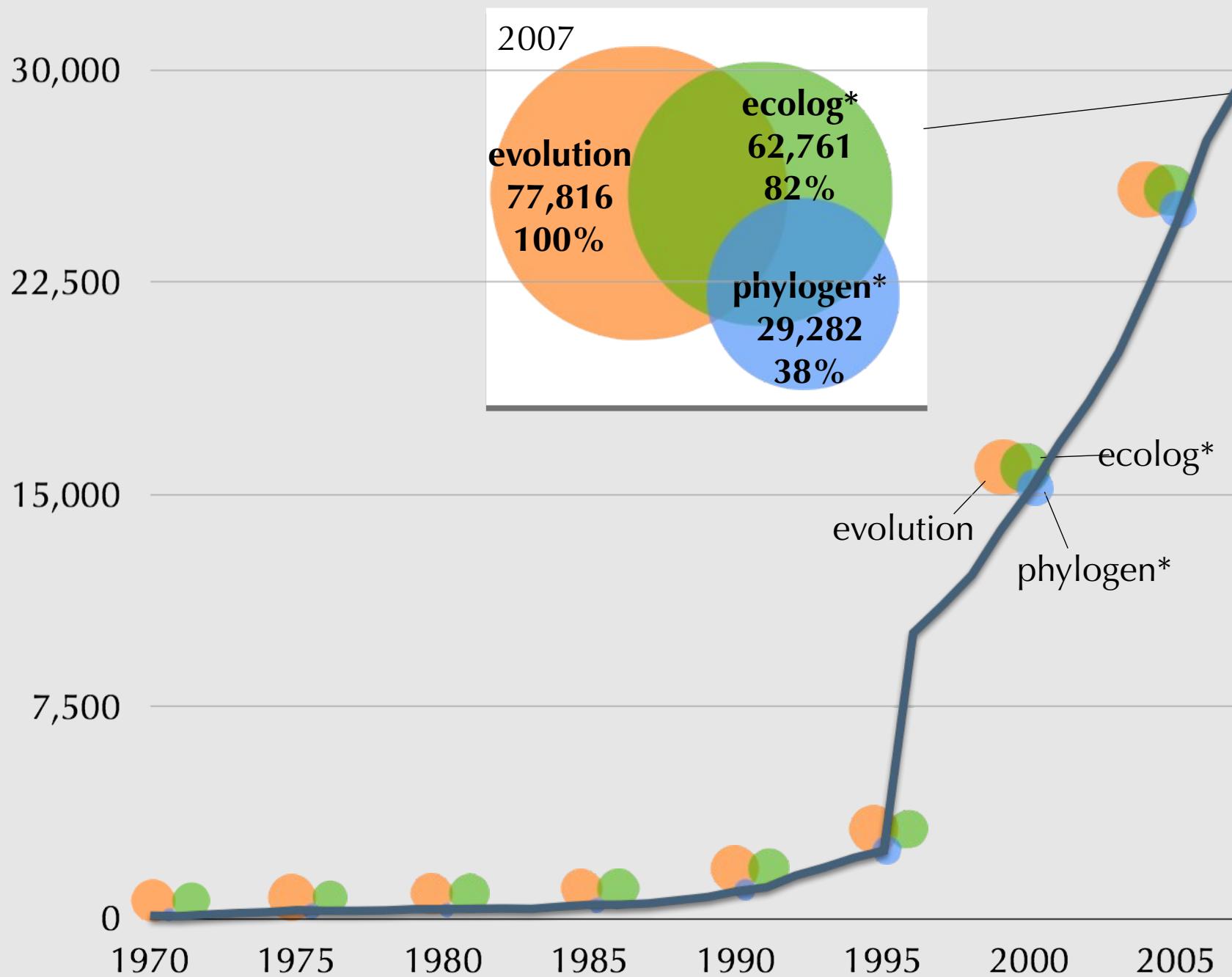


Figure 18. The phylogenetic kinship relations between the species of a monophyletic group, represented in two different ways.

Willi Hennig,
Phylogenetic
Systematics, 1966.





Select clade & genes



Pull sequences from GenBank



Filter sequences



Test and fix orthology



Filter again



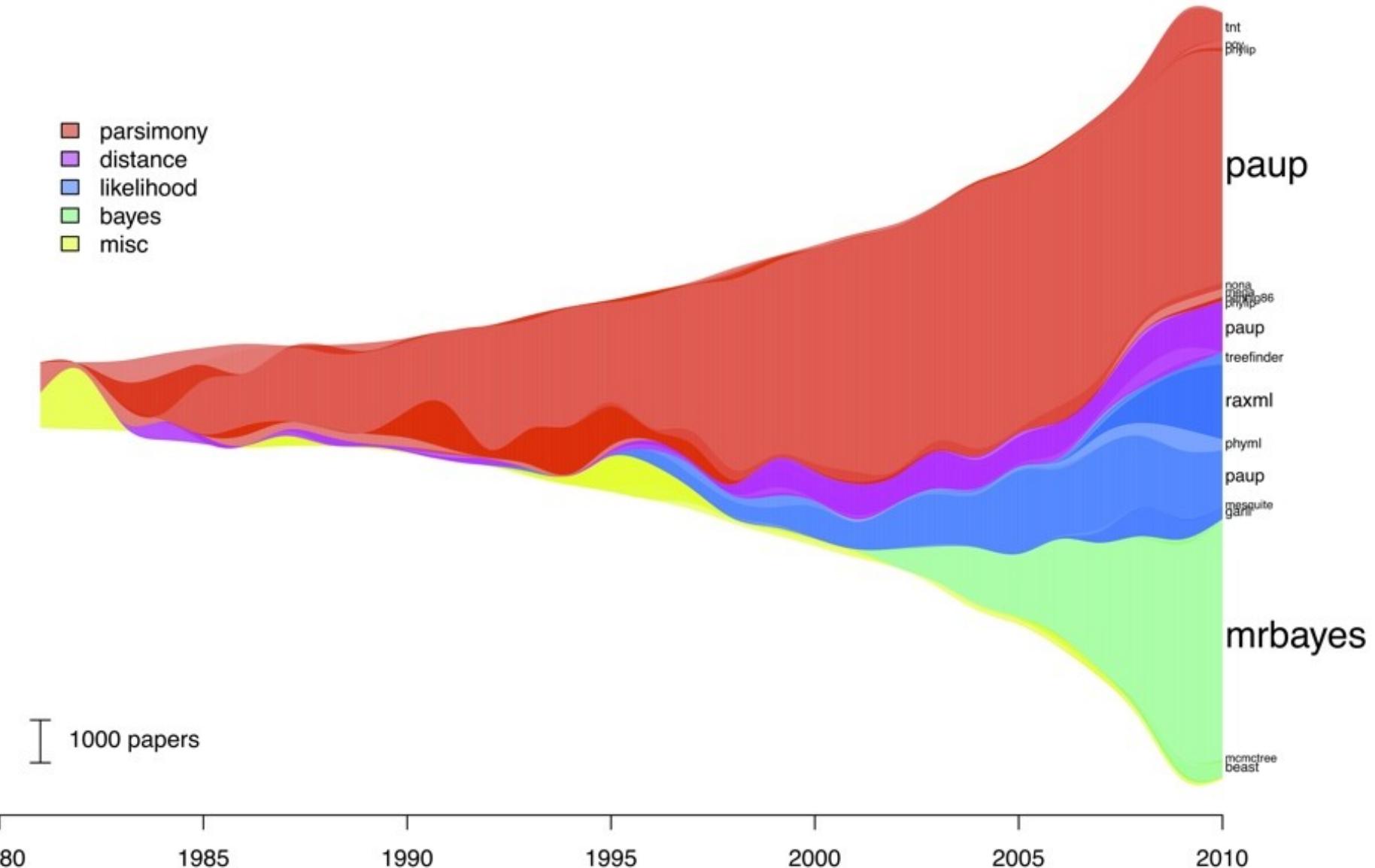
Do subalignments



Do overall alignment



Infer phylogeny



MY HOBBY:

EMBEDDING NP-COMPLETE PROBLEMS IN RESTAURANT ORDERS

CHOTCHKIES RESTAURANT	
~~ APPETIZERS ~~	
MIXED FRUIT	2.15
FRENCH FRIES	2.75
SIDE SALAD	3.35
HOT WINGS	3.55
MOZZARELLA STICKS	4.20
SAMPLER PLATE	5.80
~~ SANDWICHES ~~	
BARBECUE	6.55

WE'D LIKE EXACTLY \$15.05
WORTH OF APPETIZERS, PLEASE.

...EXACTLY? UHH...

HERE, THESE PAPERS ON THE KNAPSACK
PROBLEM MIGHT HELP YOU OUT.

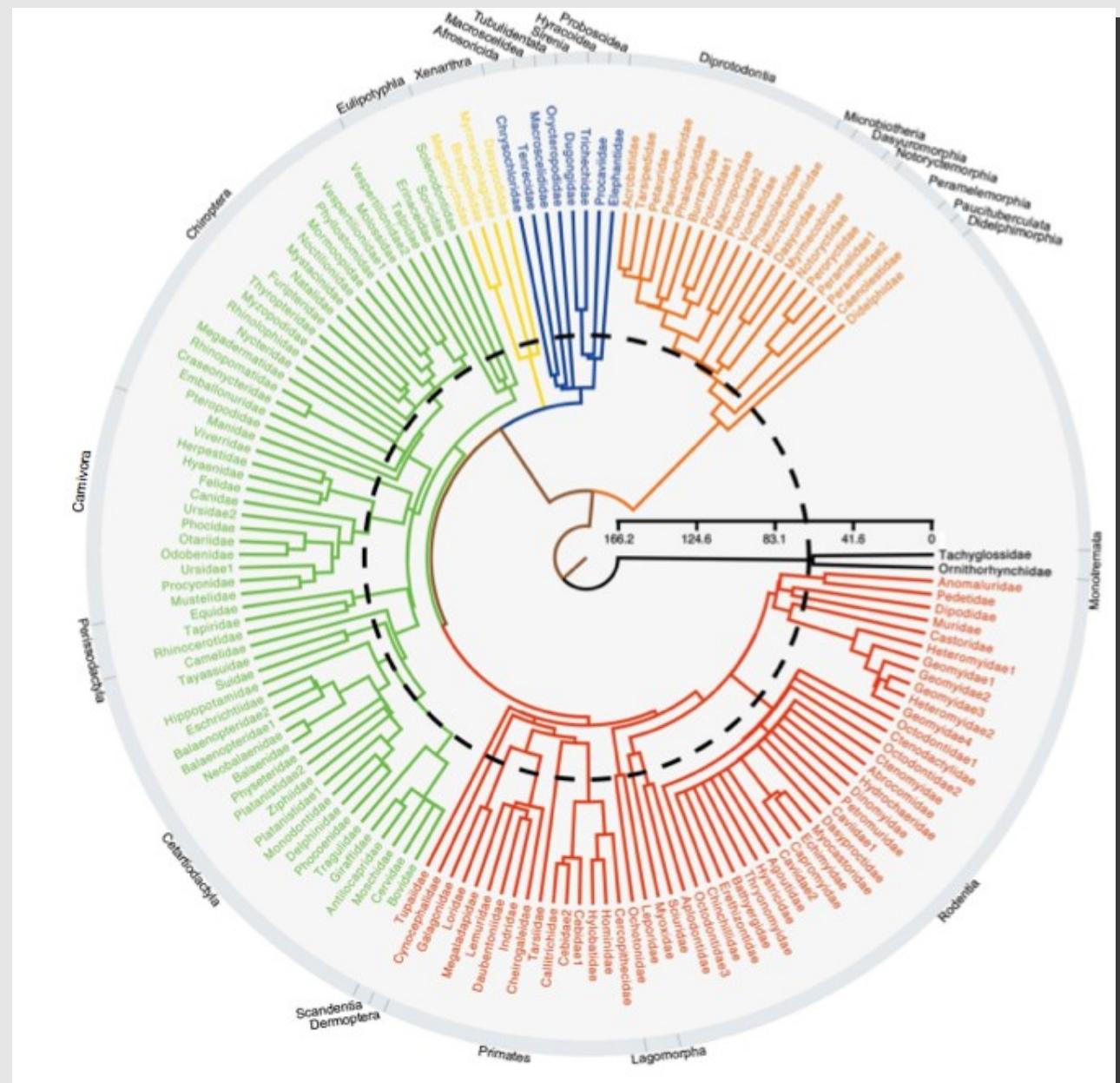
LISTEN, I HAVE SIX OTHER
TABLES TO GET TO -

-AS FAST AS POSSIBLE, OF COURSE. WANT
SOMETHING ON TRAVELING SALESMAN?

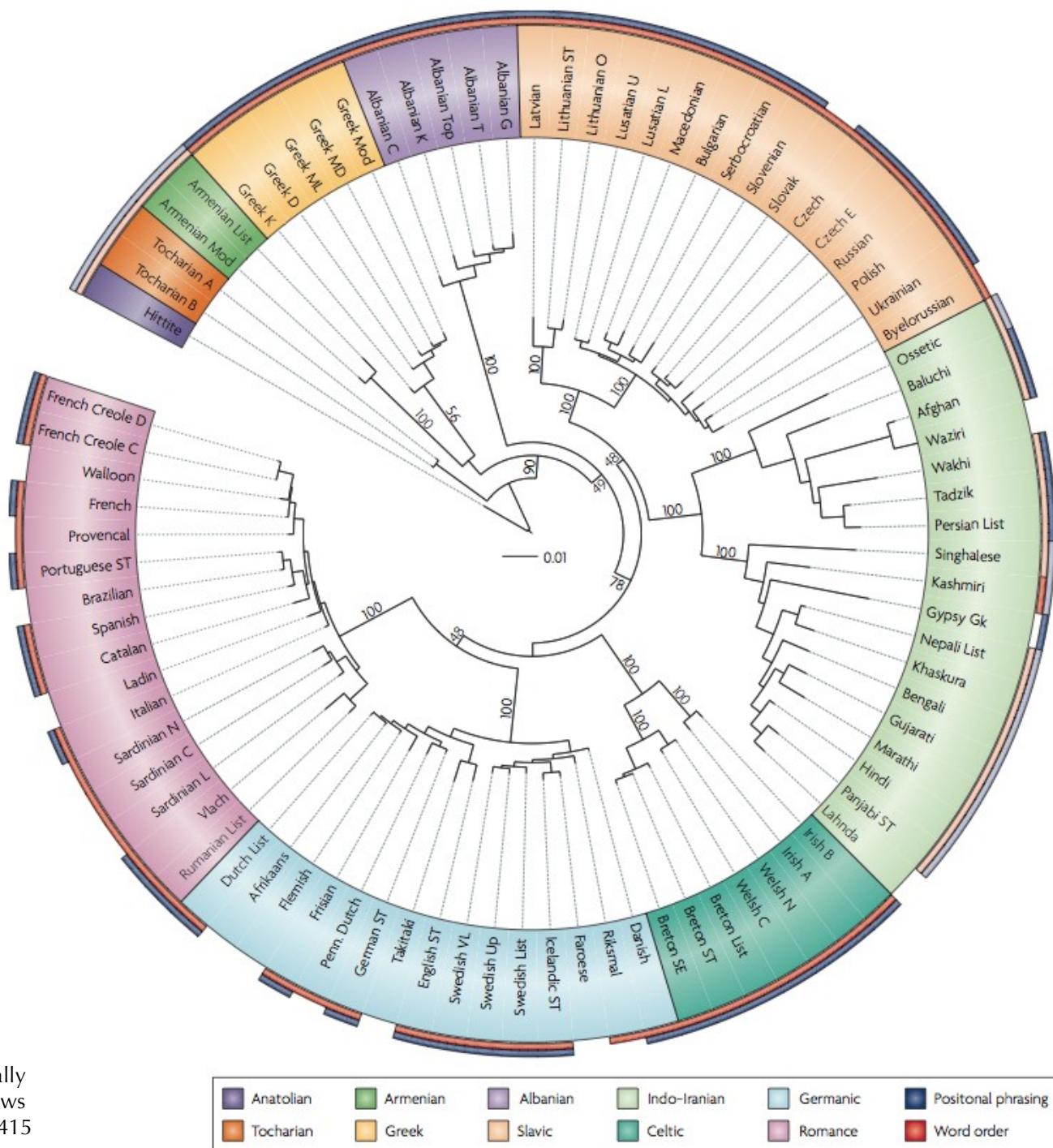


<http://xkcd.com/287/>

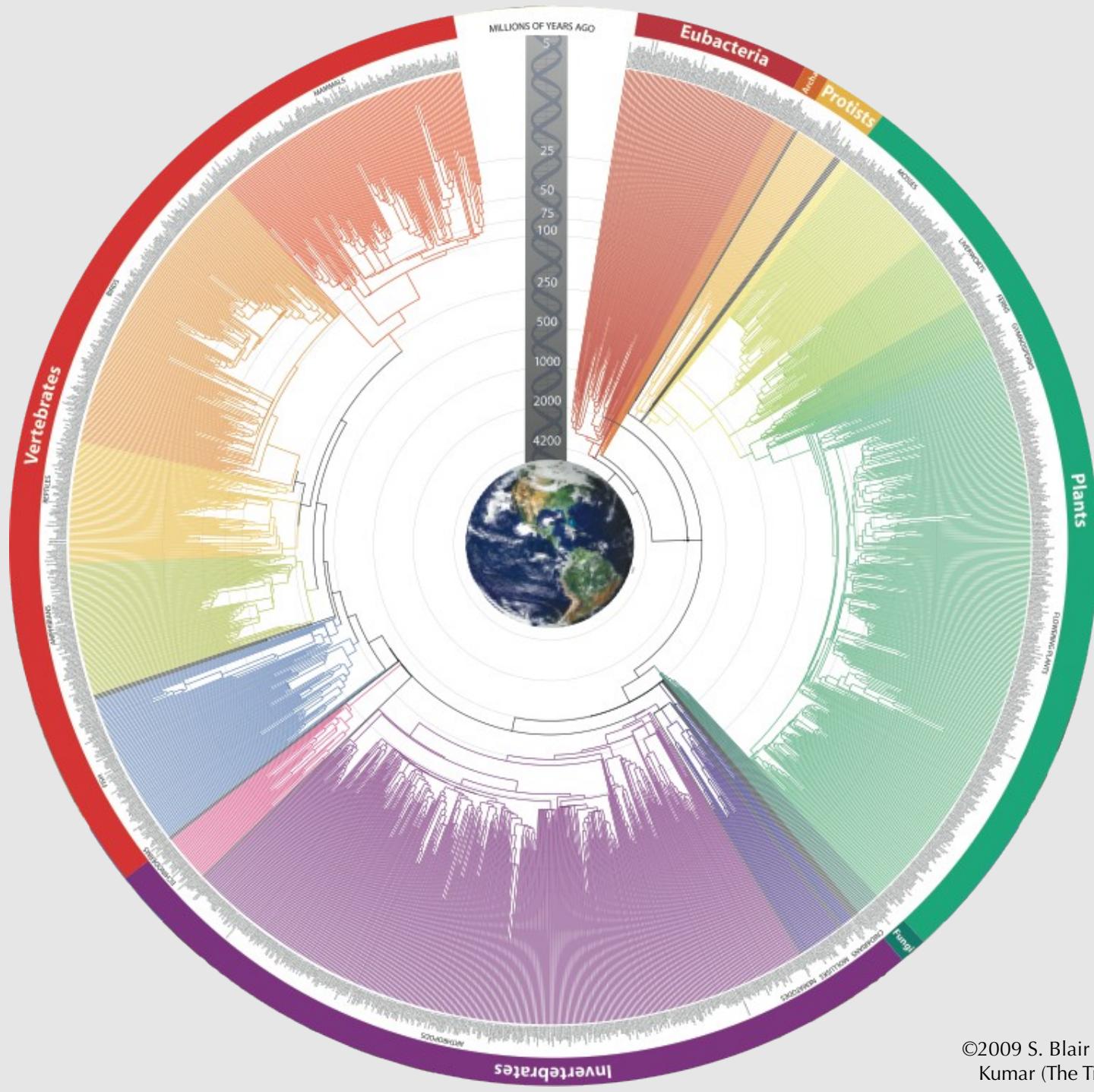
99% of mammal species
(summary tree shown)



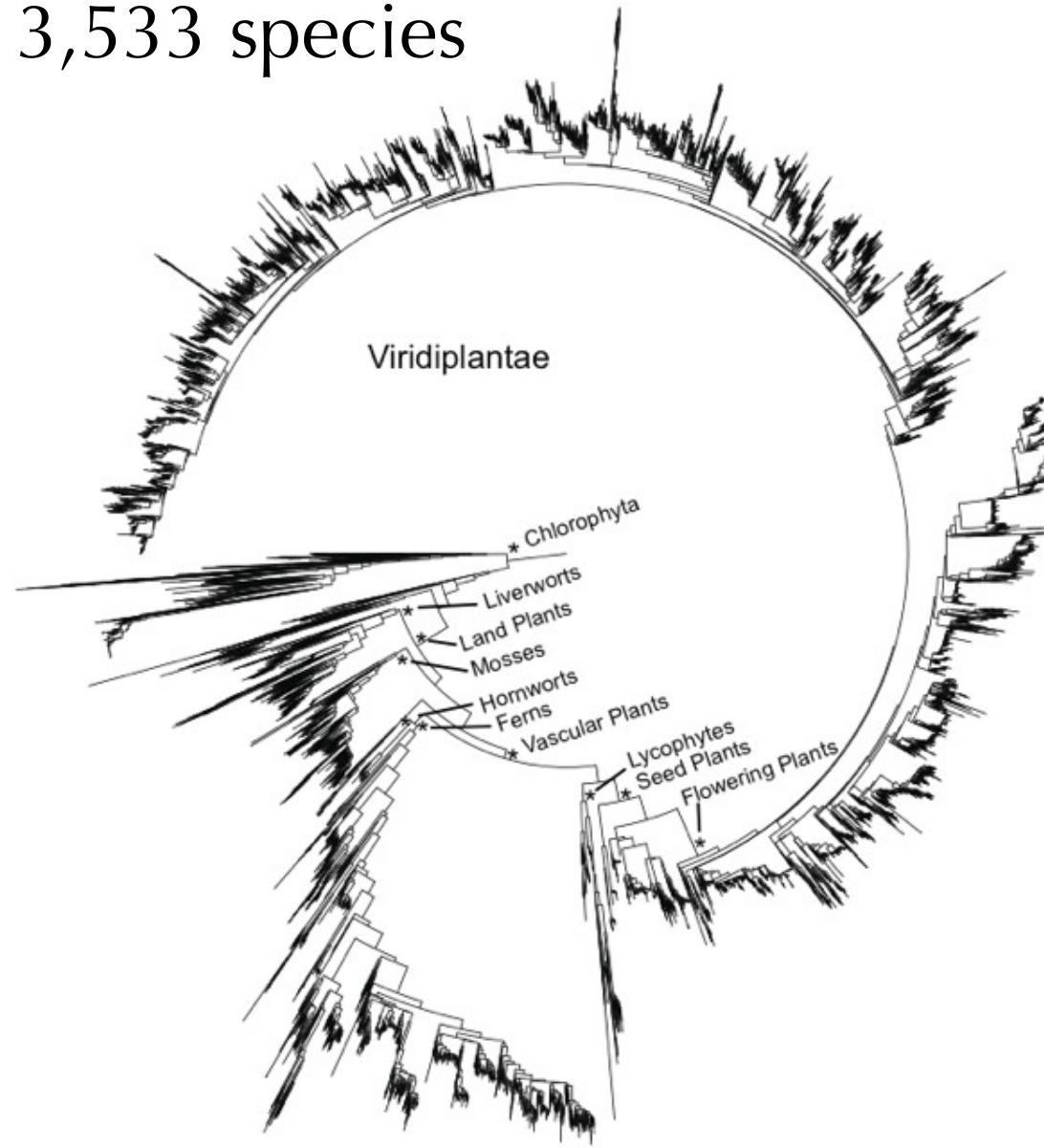
Bininda-Emonds et al. The delayed rise of present-day mammals. Nature (2007)



Pagel. Human language as a culturally transmitted replicator. Nature Reviews Genetics (2009) vol. 10 (6) pp. 405-415



13,533 species



Needed 32
GB of RAM
to run

Smith et
al. 2009

Figure 3

Maximum-likelihood phylogeny for 13,533 species of green plants based on *rbcL* DNA sequences. The data matrix was constructed using the mega-phylogeny method; major clades are labeled and denoted with a star.

1E+110

1E+100

1E+90

1E+80

1E+70

1E+60

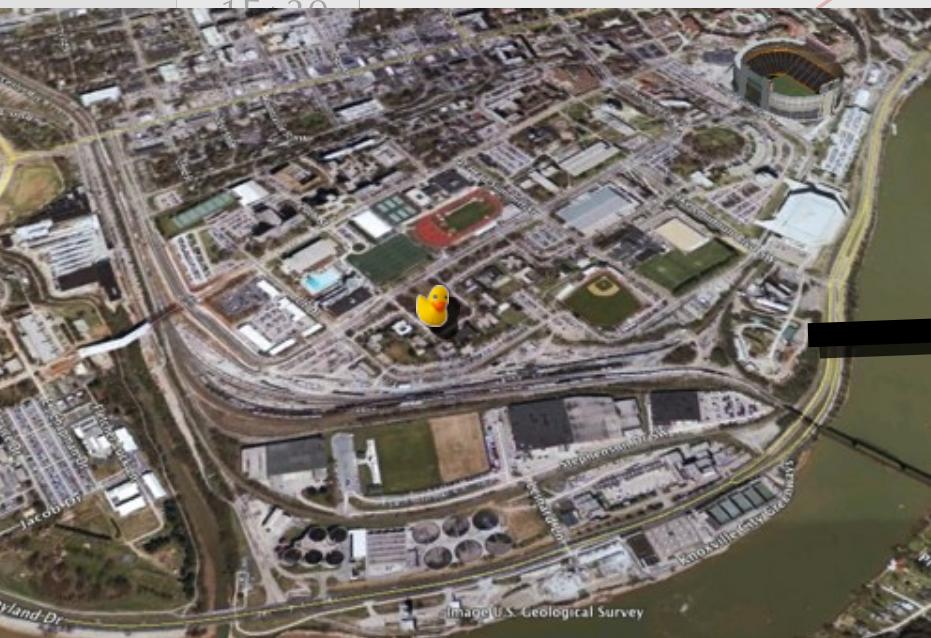
1E+50

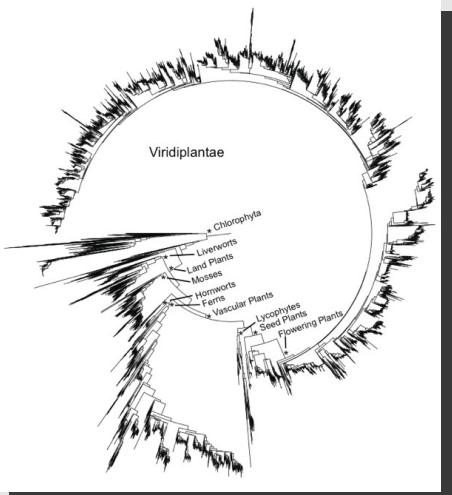
1E+40

1E+30

Number of atoms in the universe

2^N
trees
 10^N

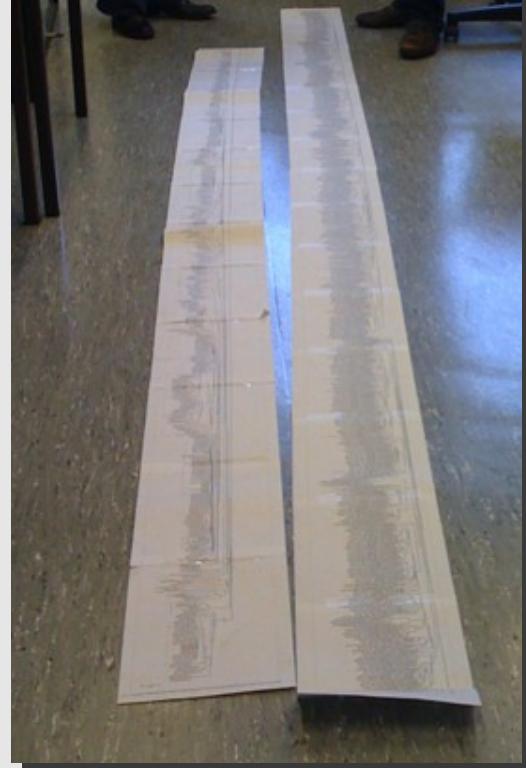




13,533 names

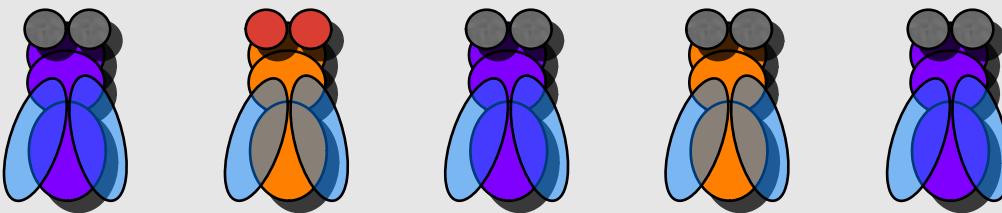


largest computer monitors:
3280×2048 (can be tiled)



Laser printer: effectively
3600 × 4725 (can be tiled)

5 fly species

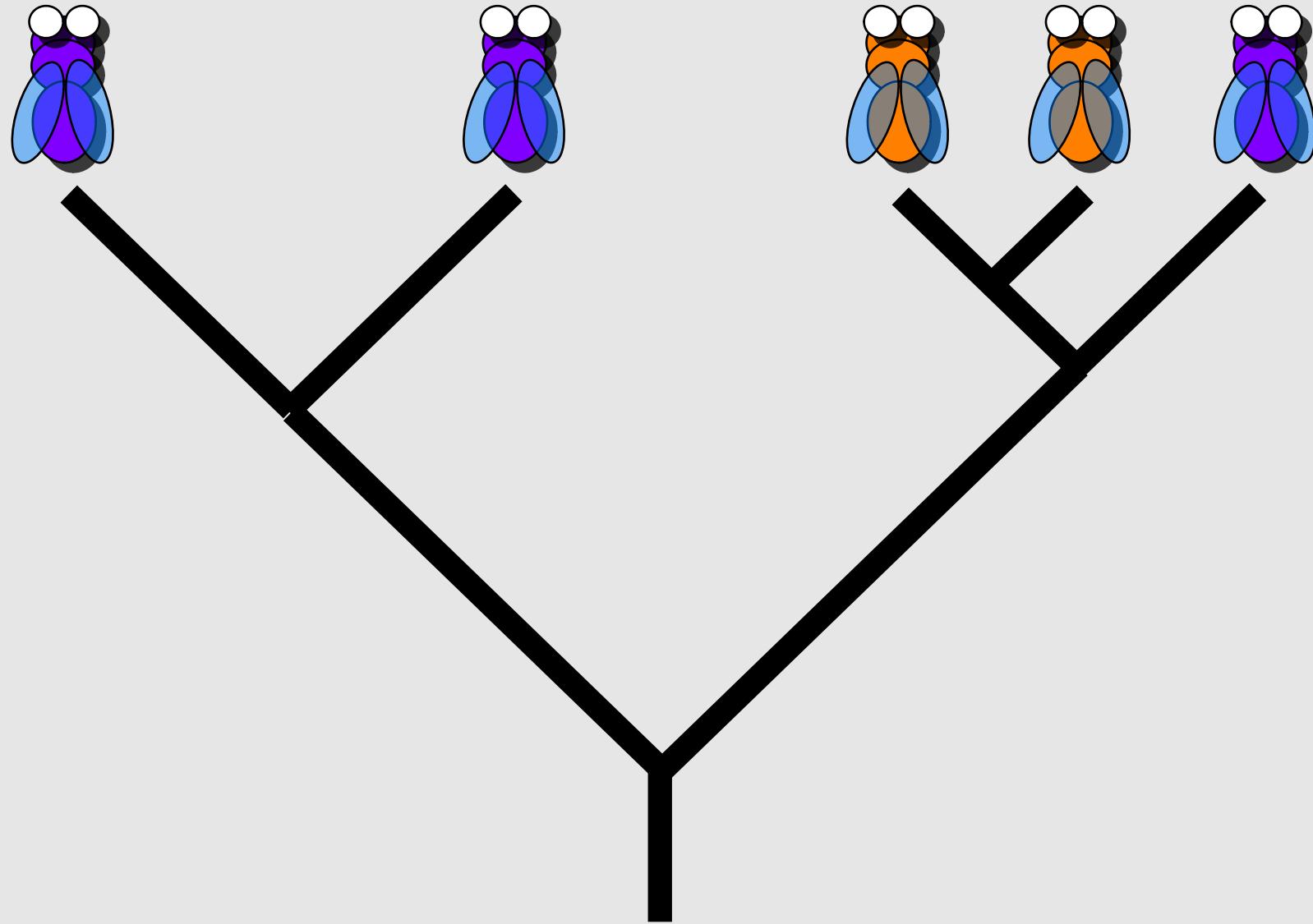


Differ in body color and eye color

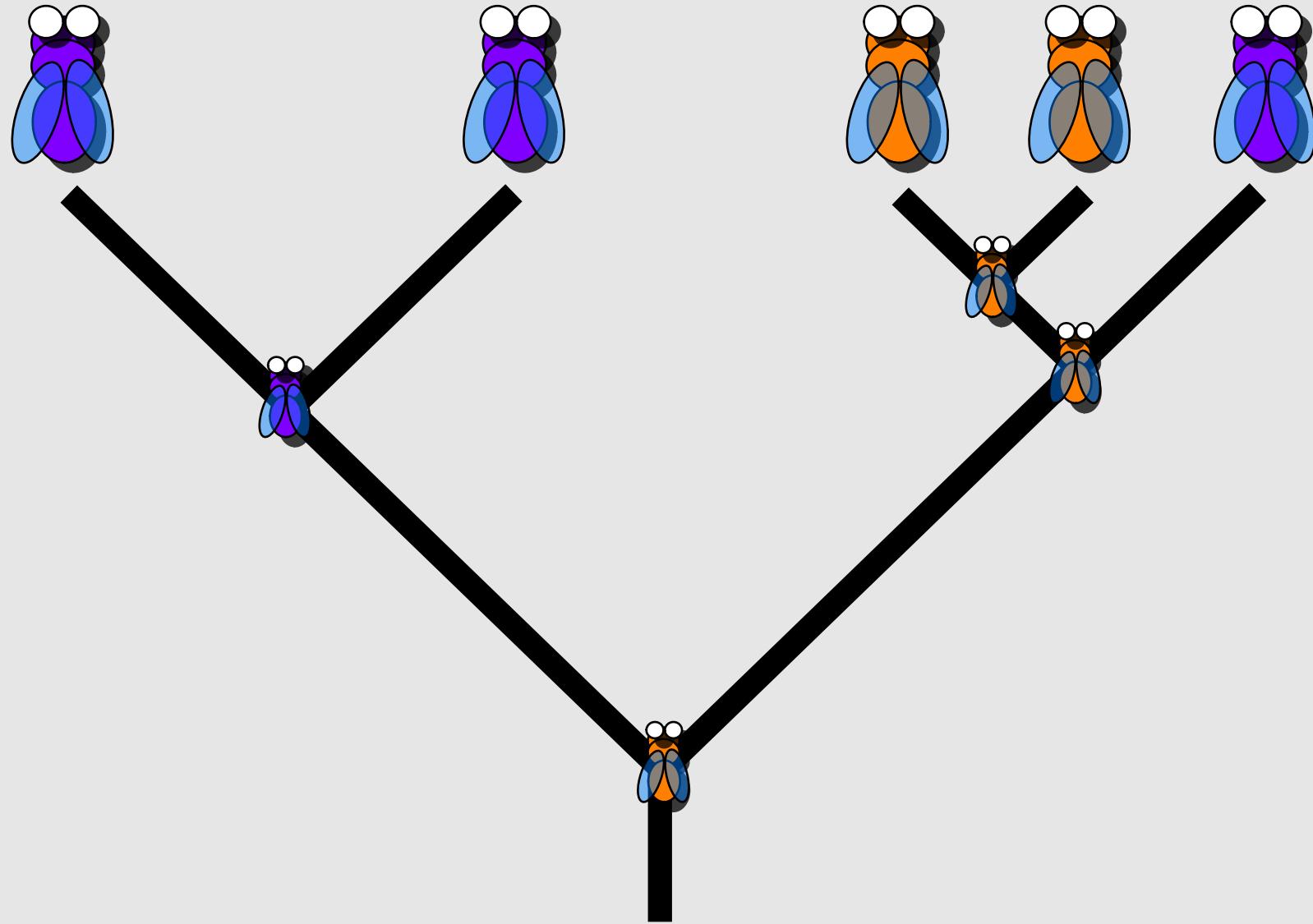
What is history of trait evolution?

The orange ones feed on oranges. Did they become orange to better feed on oranges (camouflage), or were they already that color?

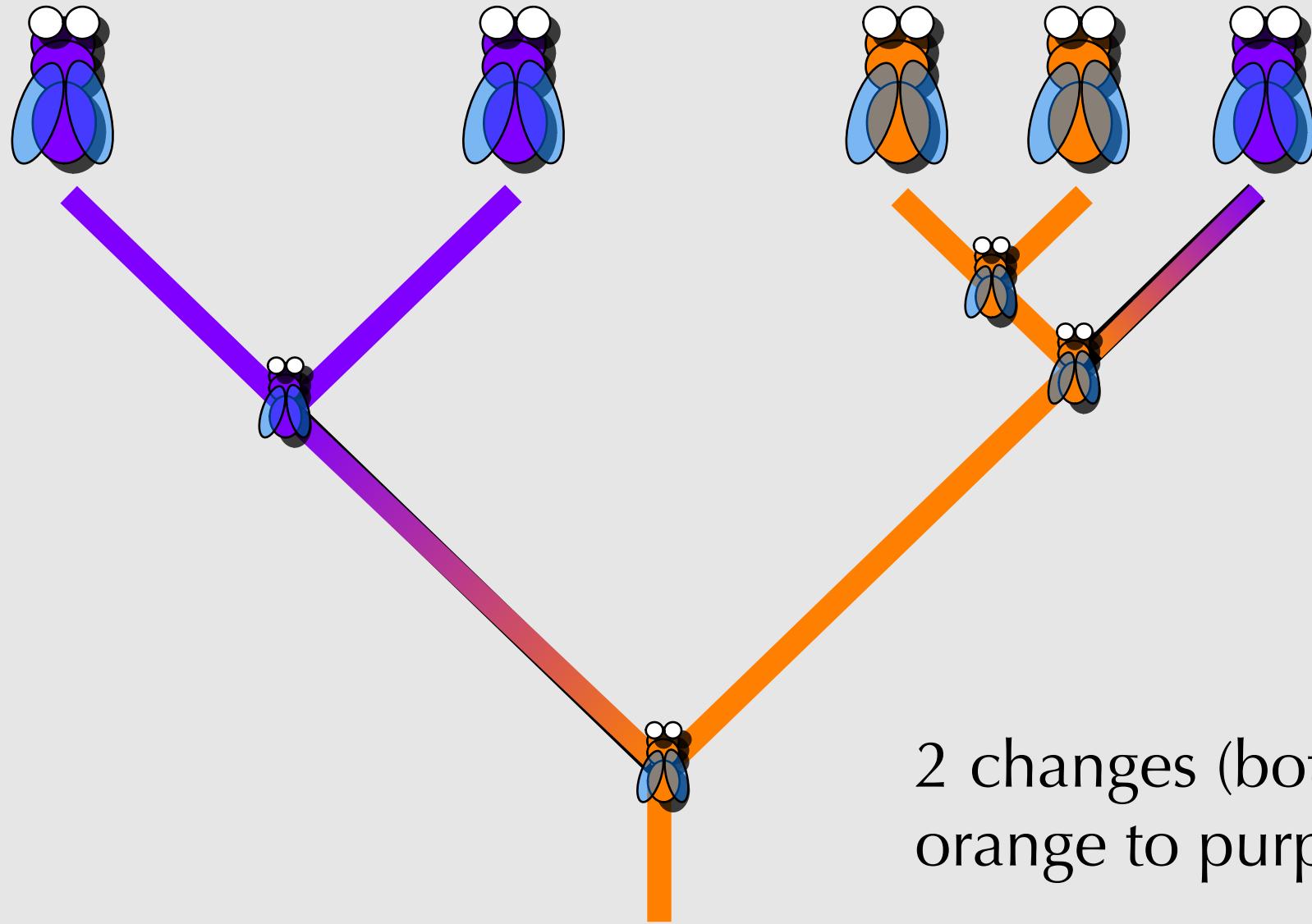
Simple reconstruction: parsimony. Minimize changes.
(look at body color alone for example)



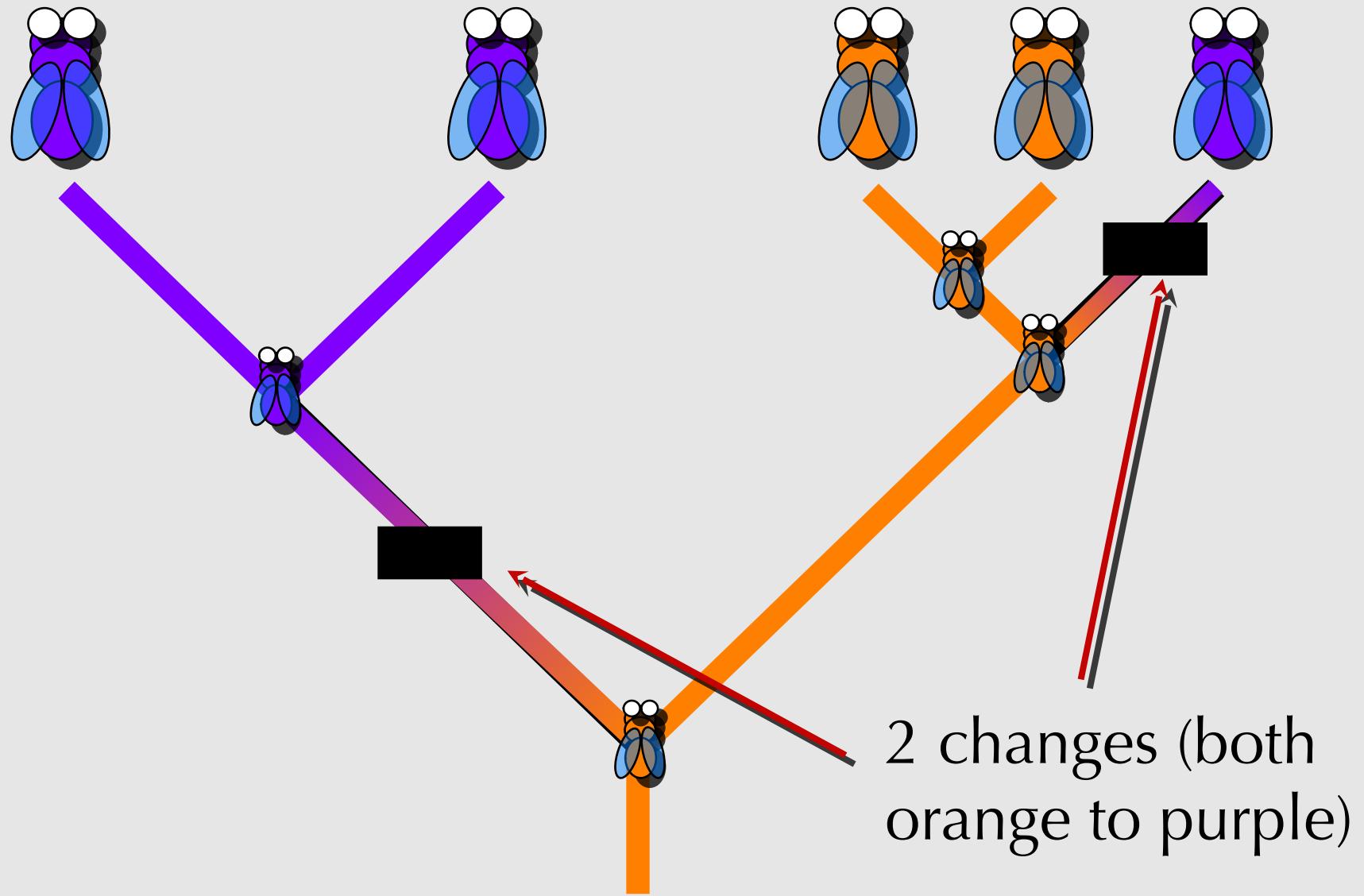
Simple reconstruction: parsimony. Minimize changes.
One possible reconstruction:



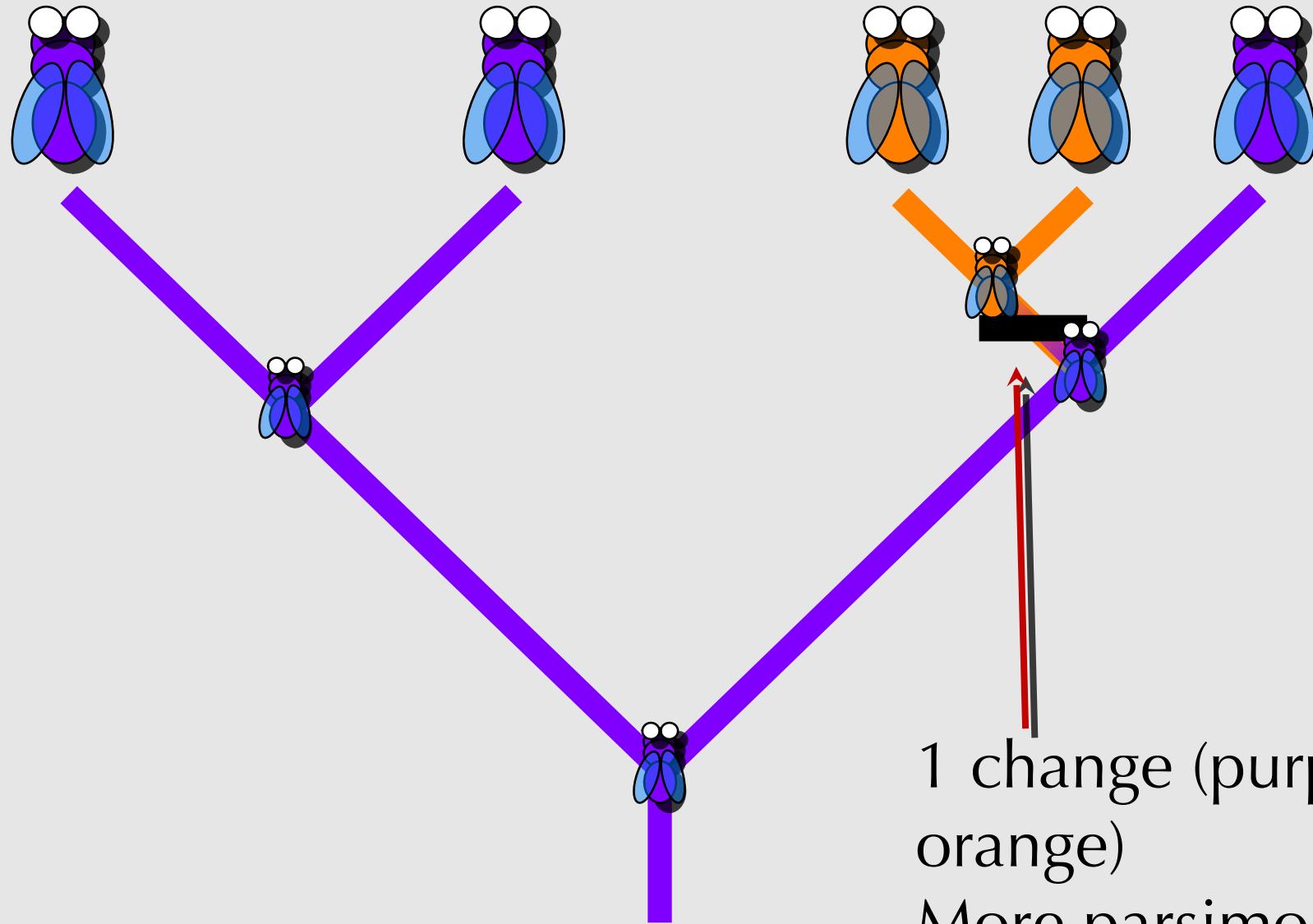
Simple reconstruction: parsimony. Minimize changes.
One possible reconstruction:



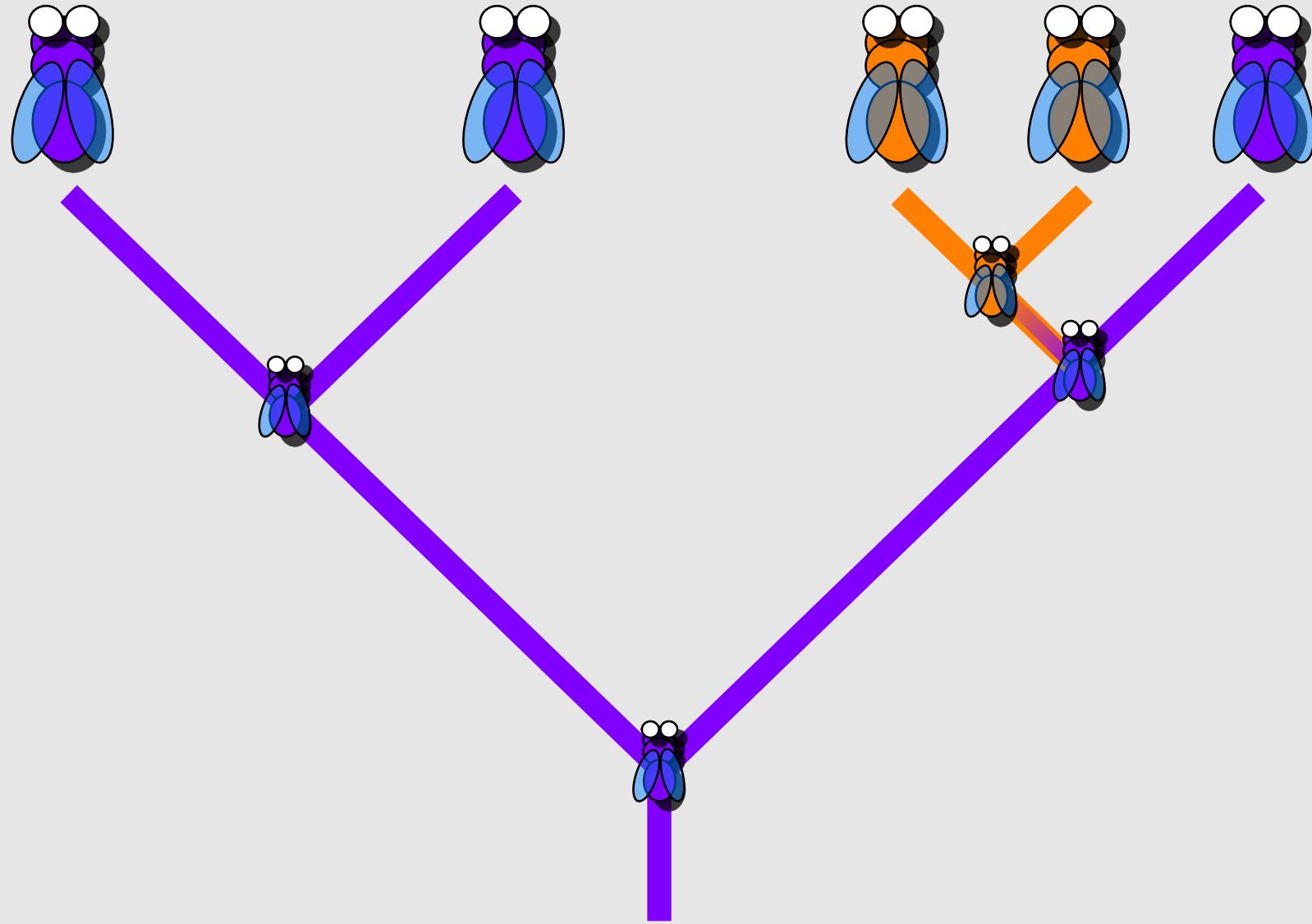
Simple reconstruction: parsimony. Minimize changes.
One possible reconstruction:



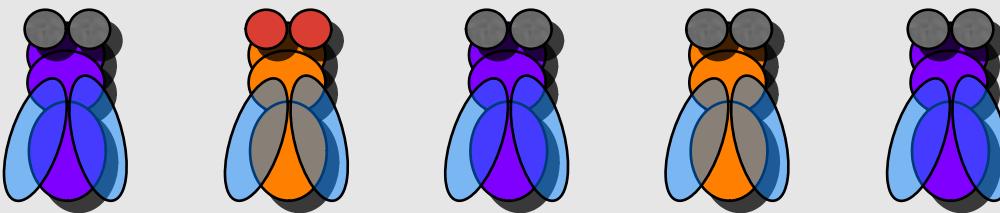
Simple reconstruction: parsimony. Minimize changes.
A different reconstruction:



What is history of trait evolution?



5 fly species

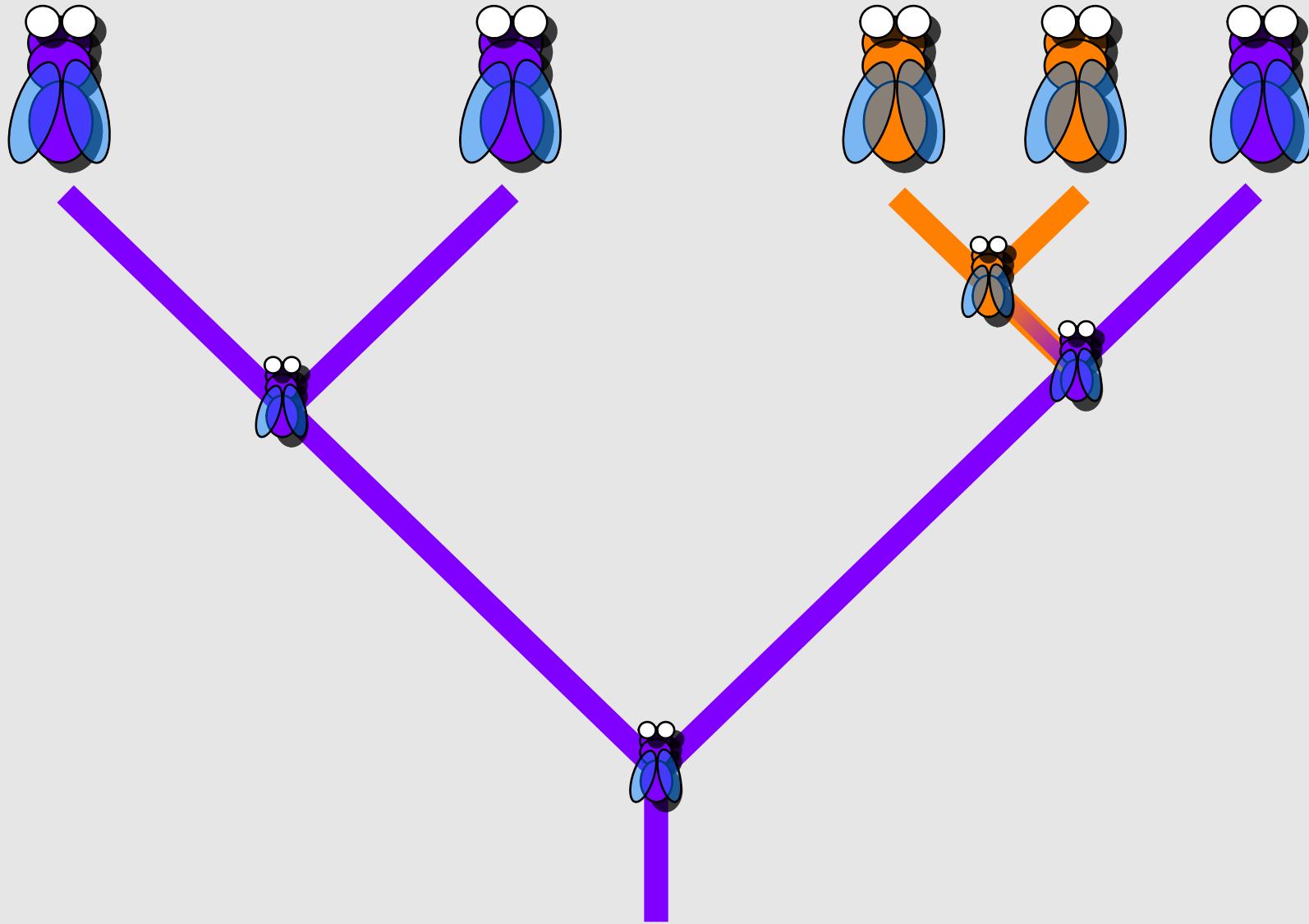


Differ in body color and eye color

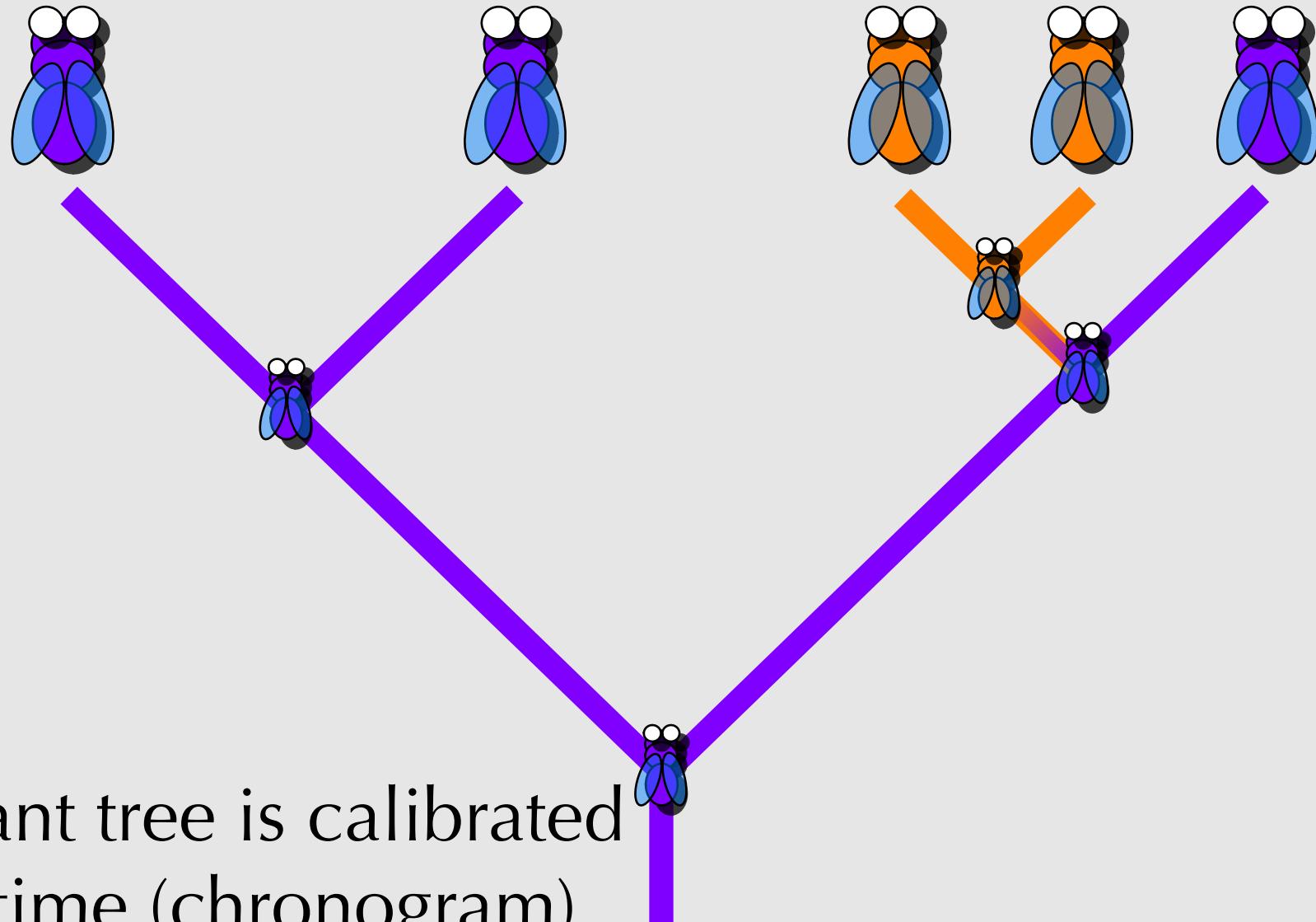
What is history of trait evolution?

The orange ones feed on oranges. Did they become orange to better feed on oranges (camouflage), or were they already that color?

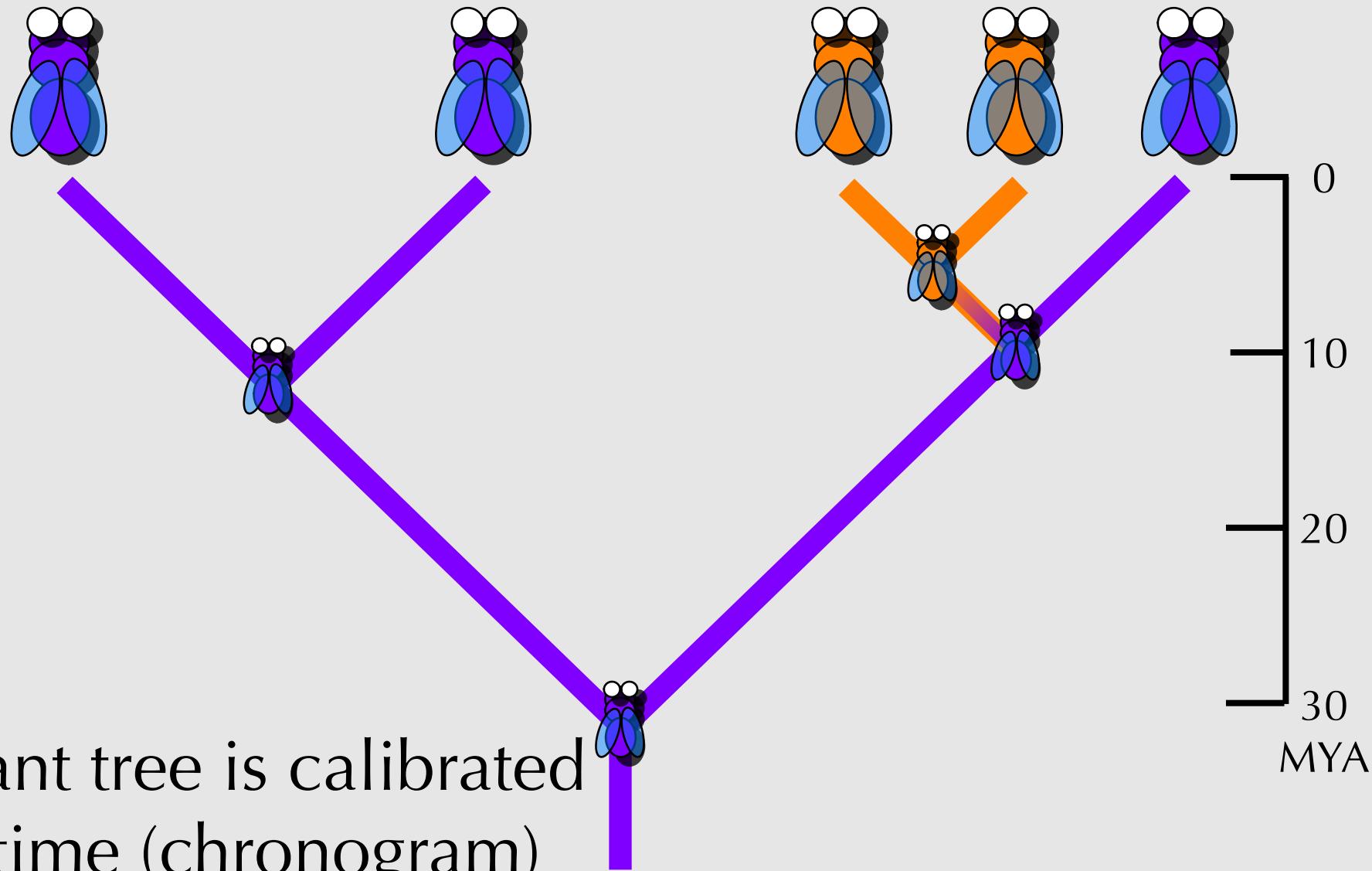
Did they become orange to better feed on oranges, or were they already that color?



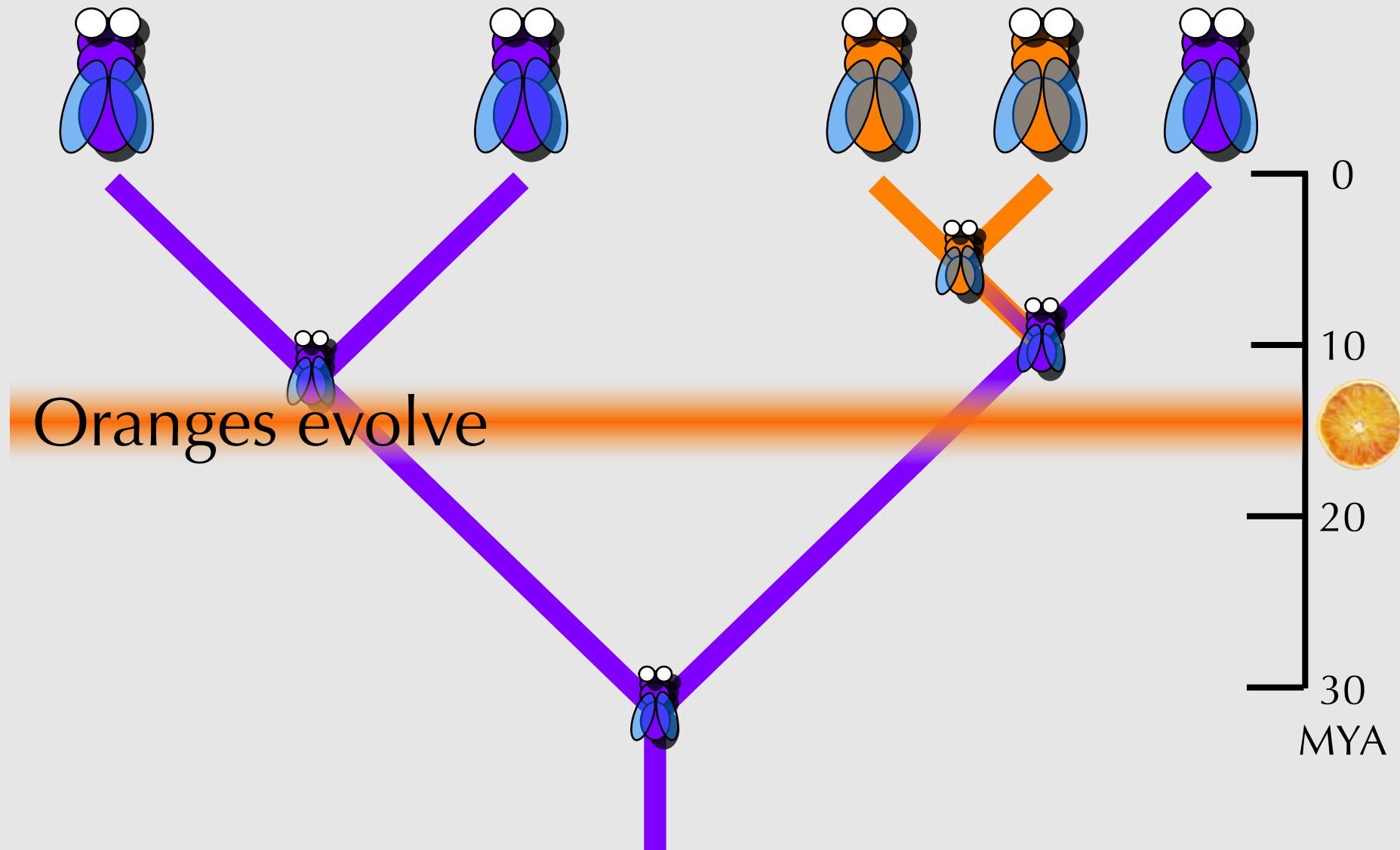
Did they become orange to better feed on oranges, or were they already that color?



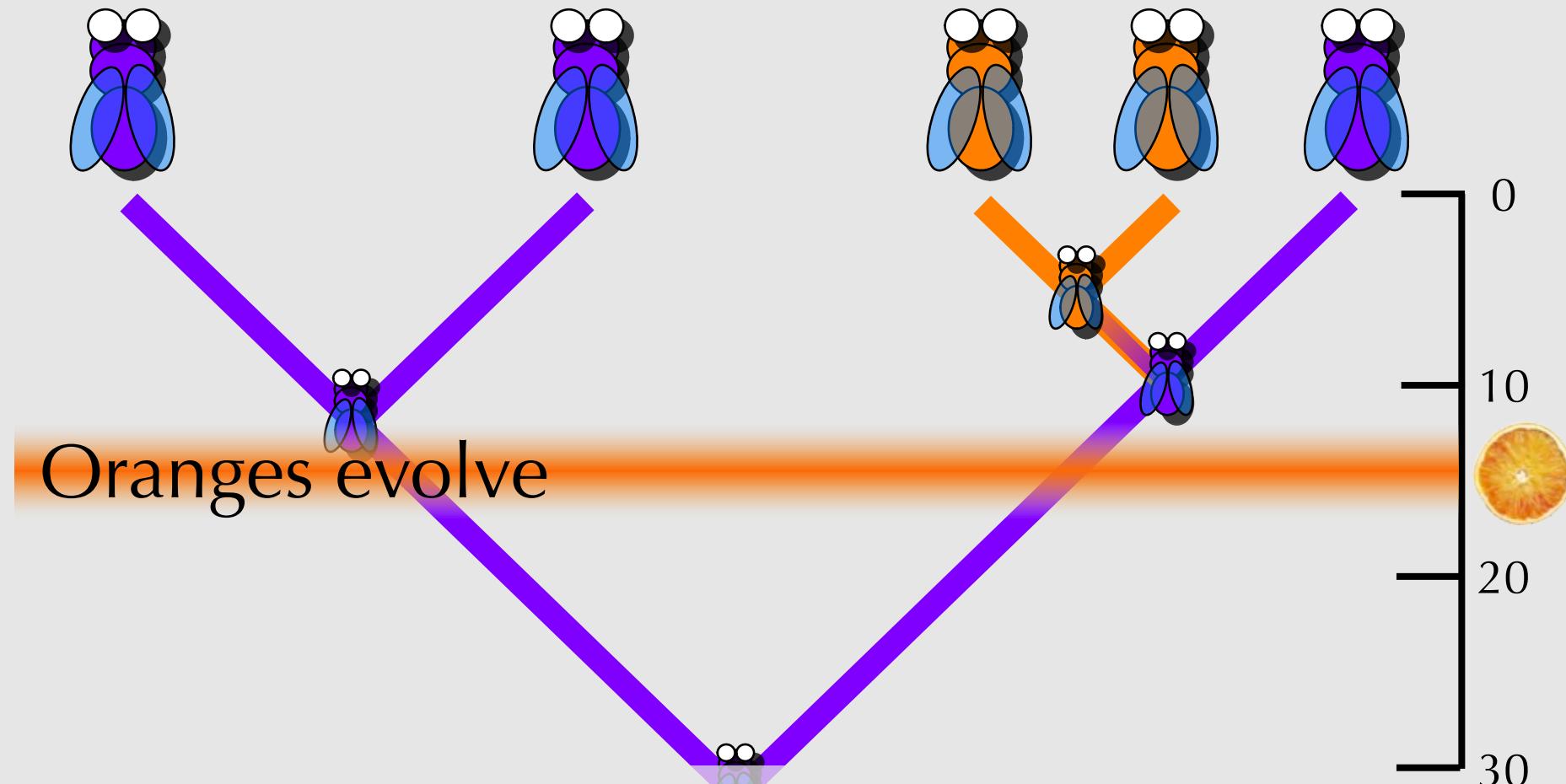
Did they become orange to better feed on oranges, or were they already that color?



Did they become orange to better feed on oranges, or were they already that color?

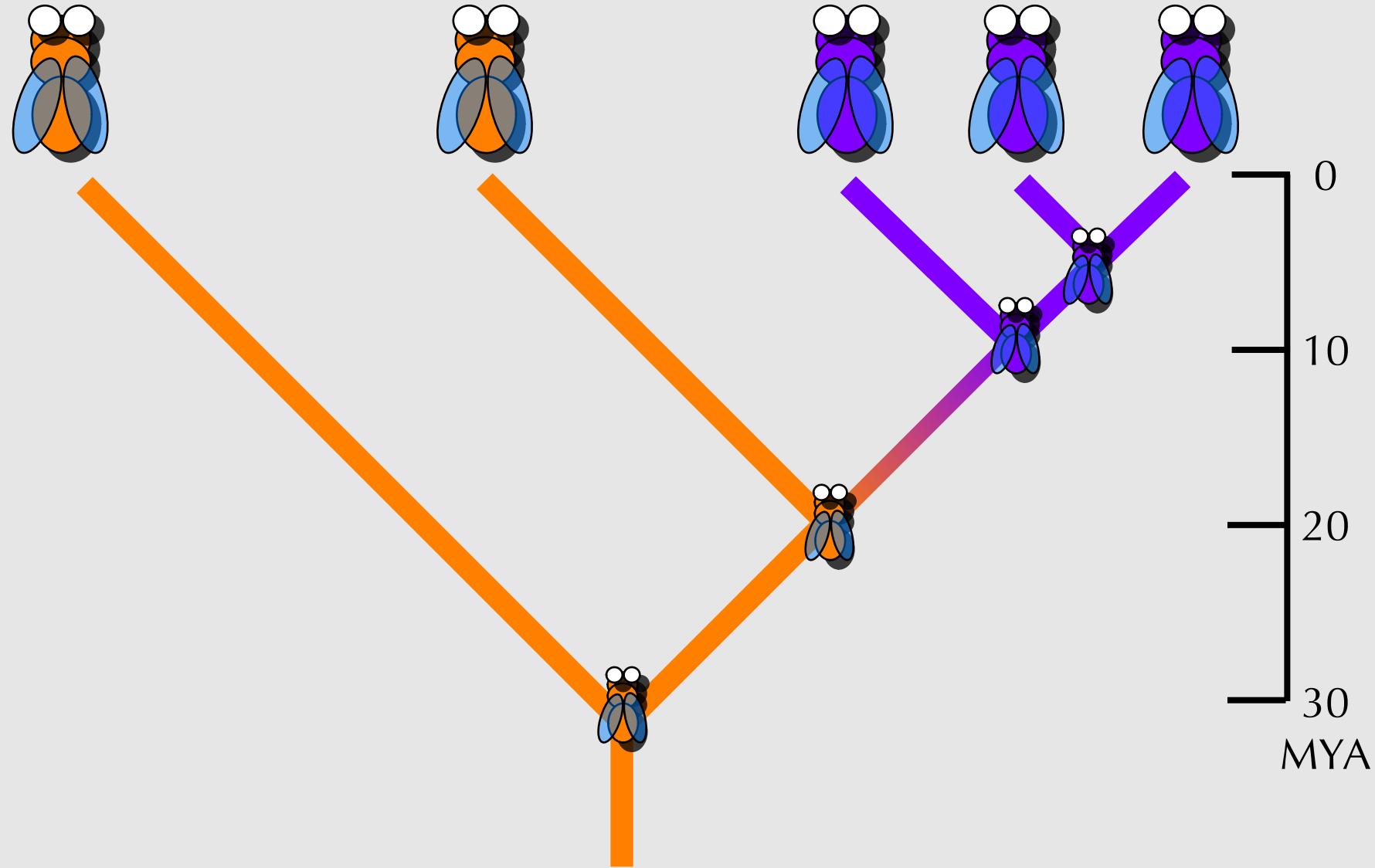


Did they become orange to better feed on oranges, or were they already that color?

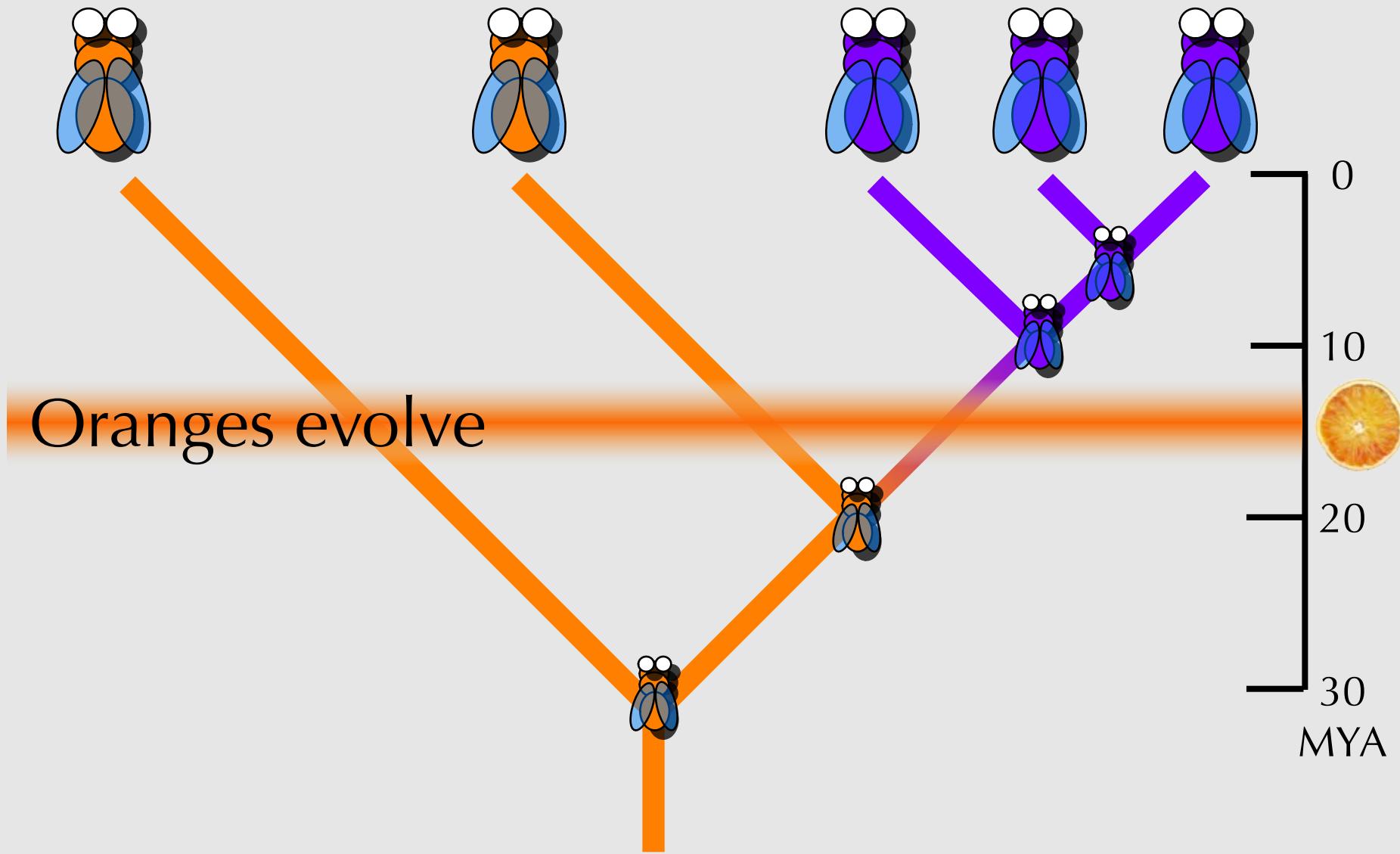


Orange color evolves AFTER orange fruit does. Consistent with hypothesis.

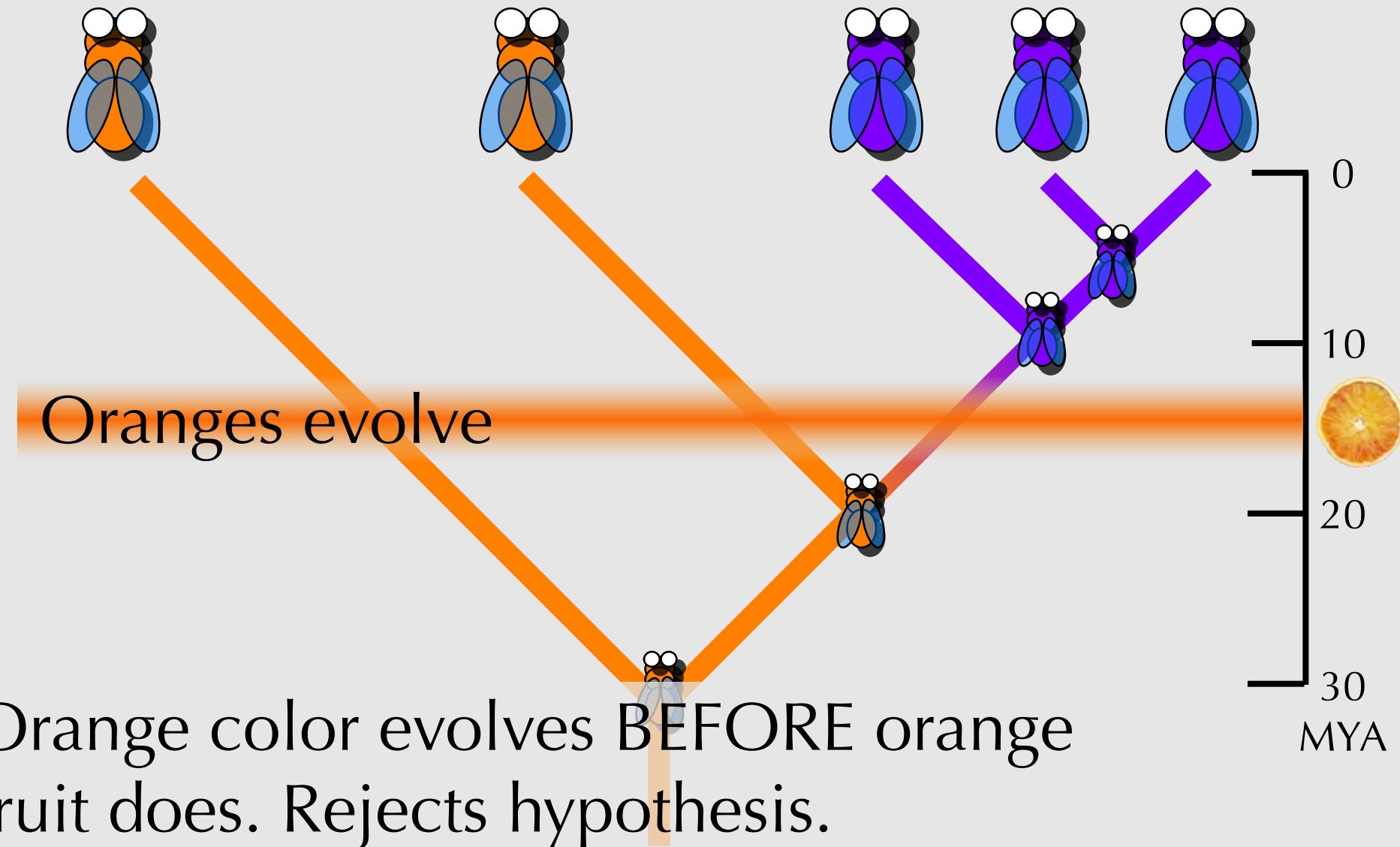
If different tree, different reconstruction



If different tree, different reconstruction

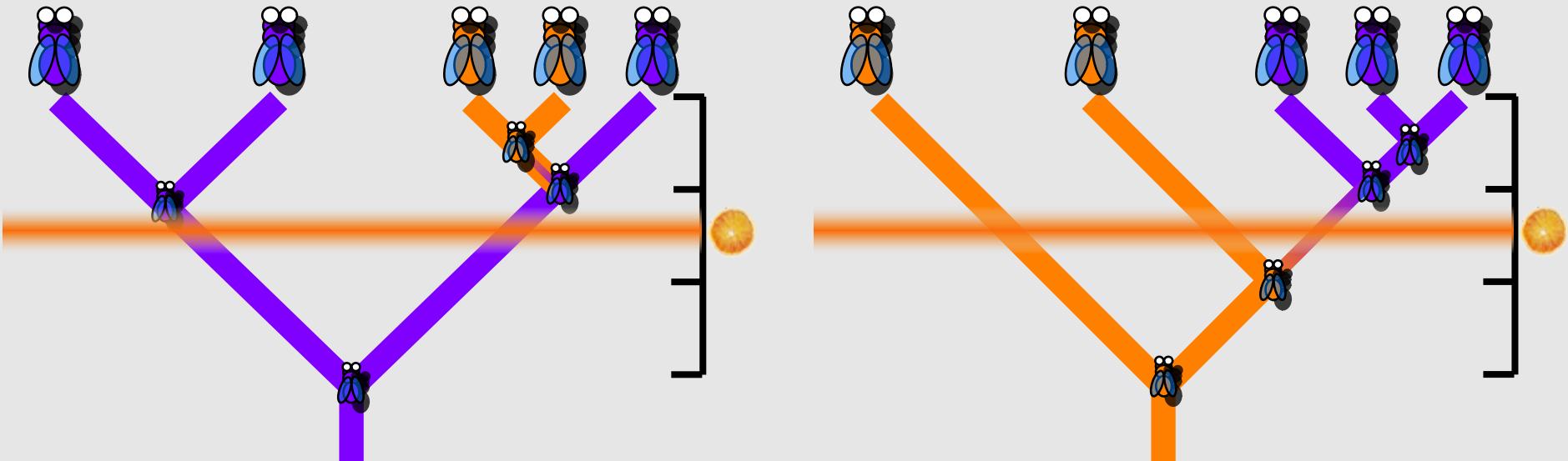


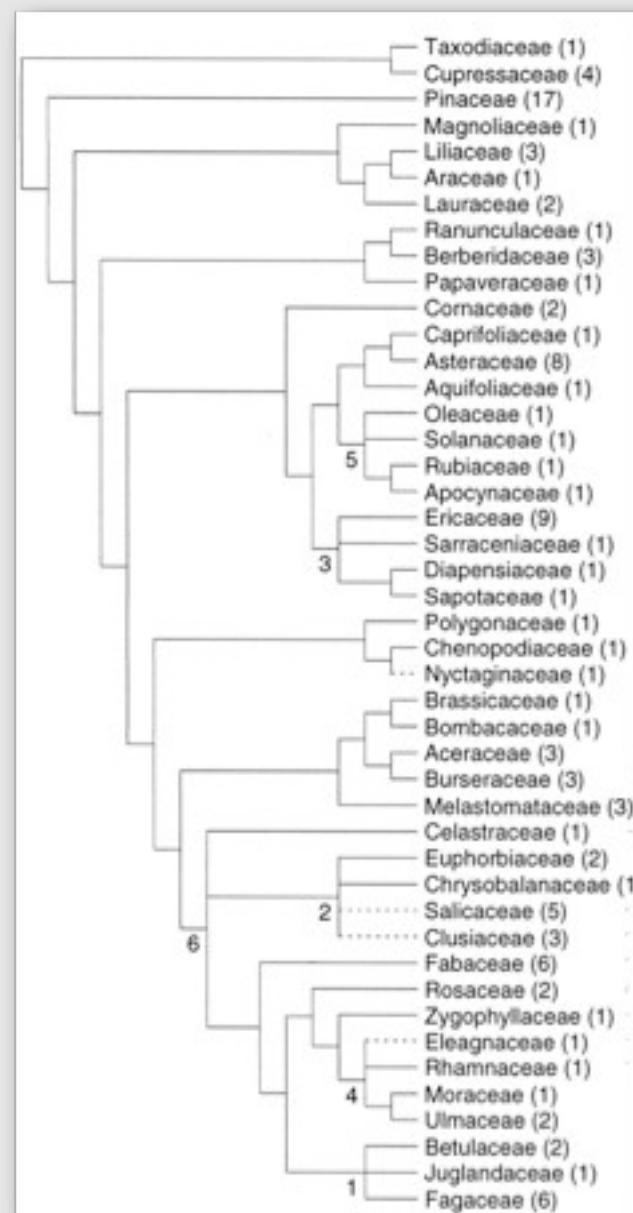
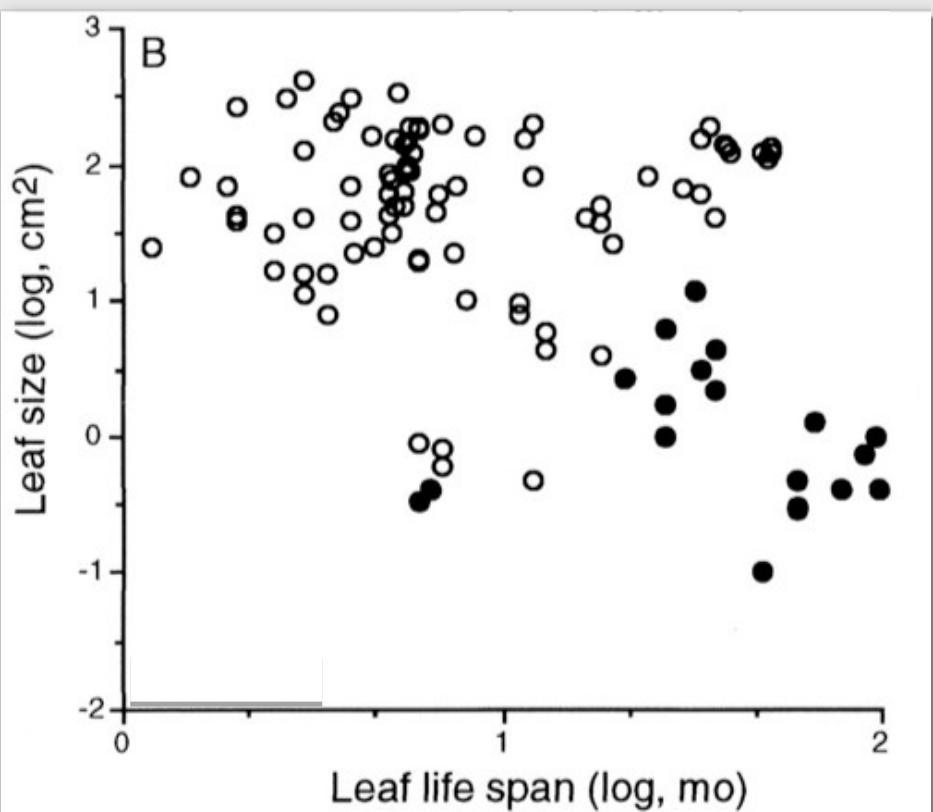
If different tree, different reconstruction



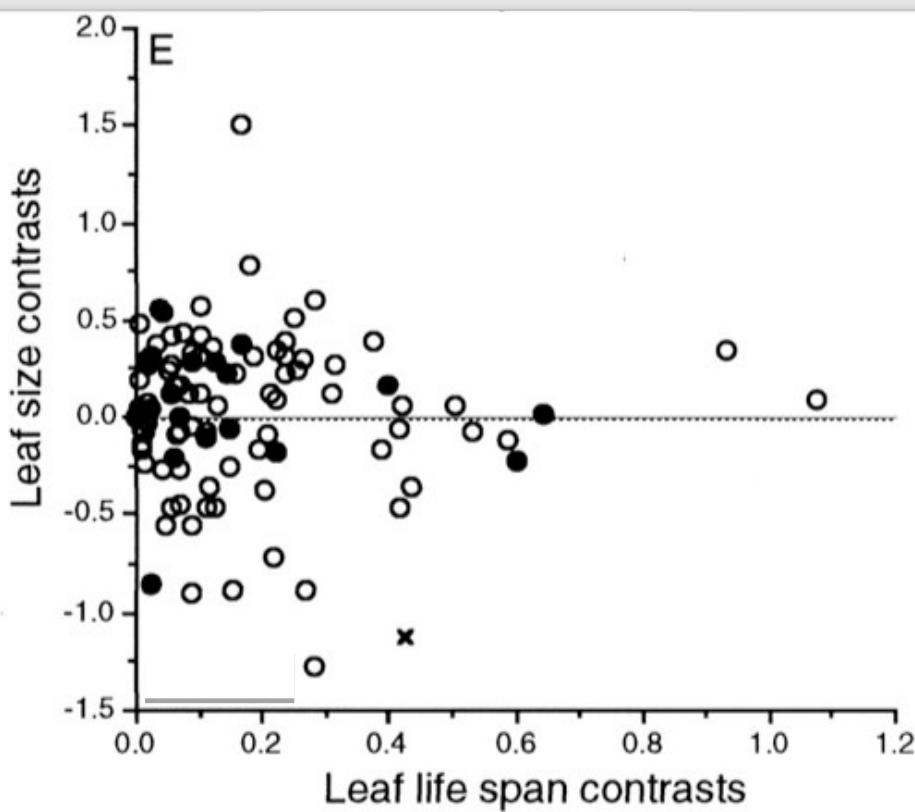
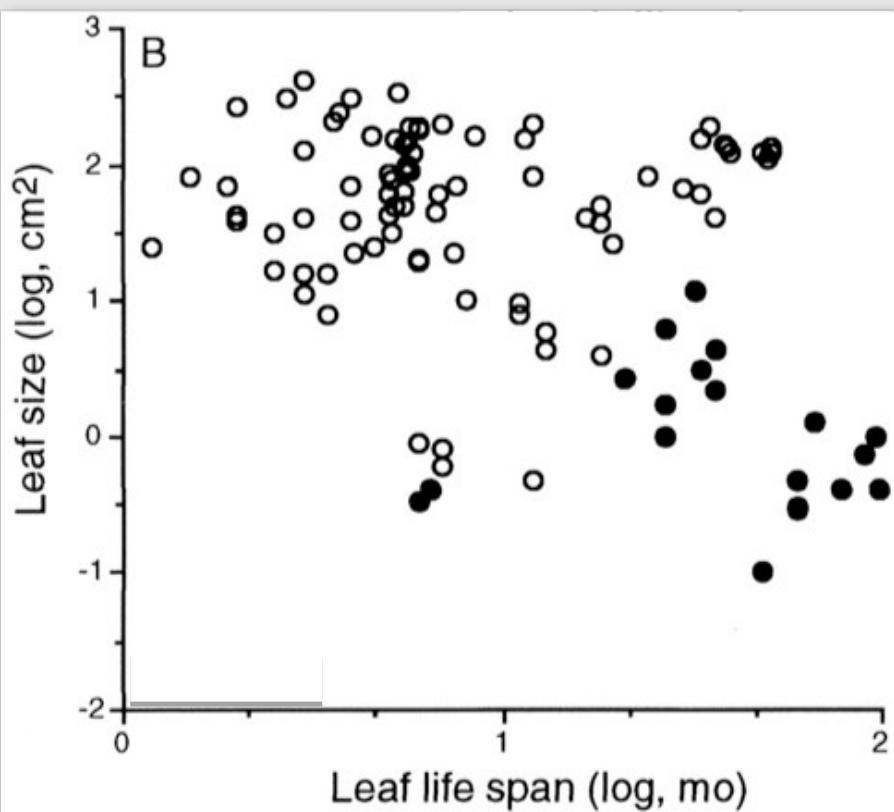
Amazing power of phylogenetics for trait evolution:

- Once you have a tree...
- And observations of extant organisms...
- You can make inferences about processes happening millions of years ago.



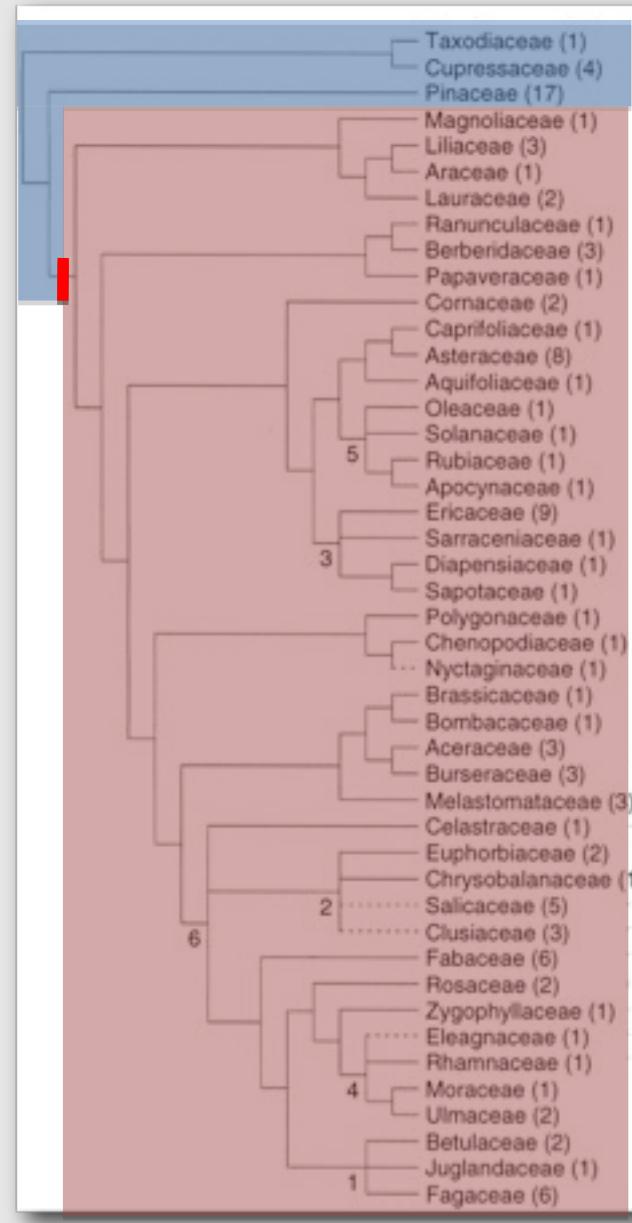


Ackerly and Reich. Convergence and correlations among leaf size and function in seed plants: a comparative test using independent contrasts. Am. J. Bot. (1999) vol. 86 (9) pp. 1272-1281

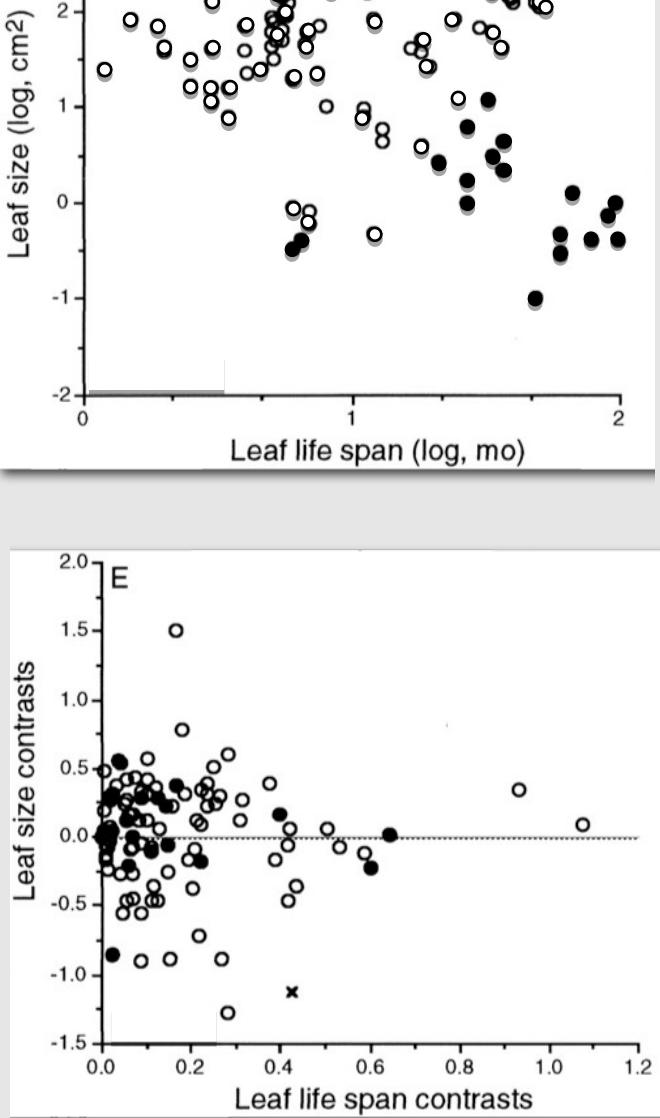


Ackerly and Reich. Convergence and correlations among leaf size and function in seed plants: a comparative test using independent contrasts. Am. J. Bot. (1999) vol. 86 (9) pp. 1272-1281

Small, long-lived



Big, short-lived



Ackerly and Reich. Convergence and correlations among leaf size and function in seed plants: a comparative test using independent contrasts. Am. J. Bot. (1999) vol. 86 (9) pp. 1272-1281

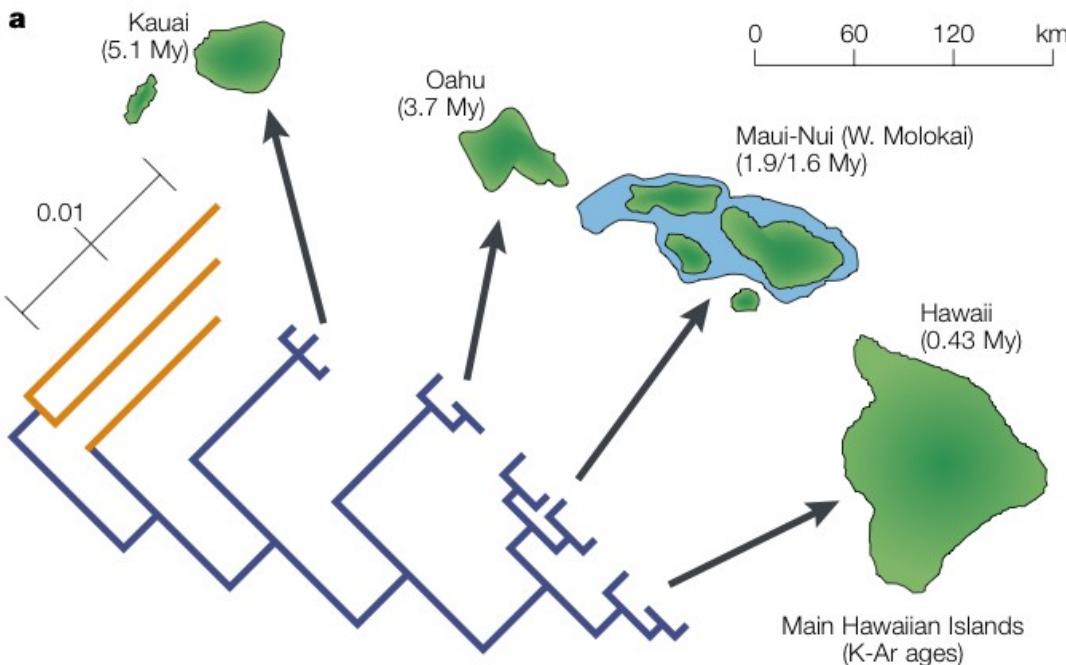
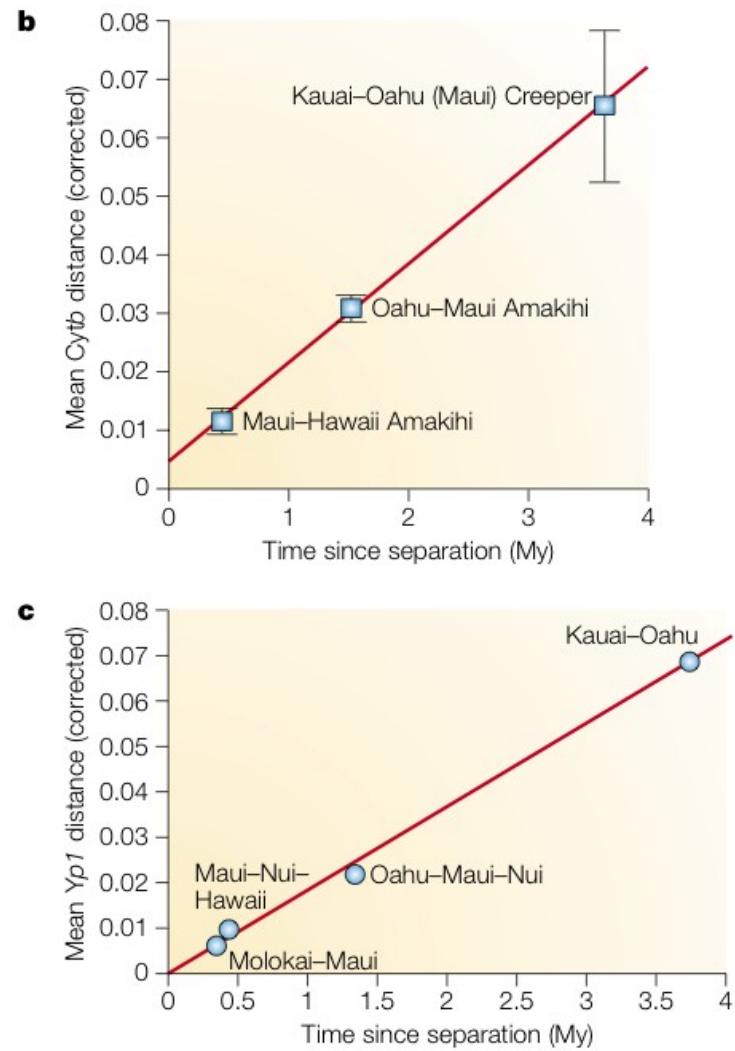


Figure 2 | A molecular clock for the Hawaiian islands. **a** | The volcanic origin of the Hawaiian islands has produced a chain of islands of increasing geological age. The phylogenetic relationships of island endemic birds (for example, the drepananine (honeycreeper) species such as the amakihi, *Hemignathus virens* and the akiapolaau *Hemignathus wilsoni*, shown in the tree) and fruitflies (*Drosophila* spp.) reflect this volcanic ‘conveyer belt’, with the species of the oldest islands forming the deepest branch of the tree, and the younger islands on the tips of the tree. Orange lines represent the outgroups. **b,c** | Molecular dates for *Hemignathus* (panel **b**) and *Drosophila* (panel **c**) confirm this order of colonization, and produce a remarkably linear relationship between genetic divergence and time when DNA distance is plotted against island age. My, million years. Figures reproduced with permission from REF. 10 © (1998) Blackwell Publishing.



Bromham and Penny. The modern molecular clock. Nature Reviews Genetics (2003) vol. 4 (3) pp. 216-224. Figure after Fleisher et al. Evolution on a volcanic conveyor belt: using phylogeographic reconstructions and K-Ar based ages of the Hawaiian islands to estimate molecular evolutionary rates. Mol Ecol (1998)

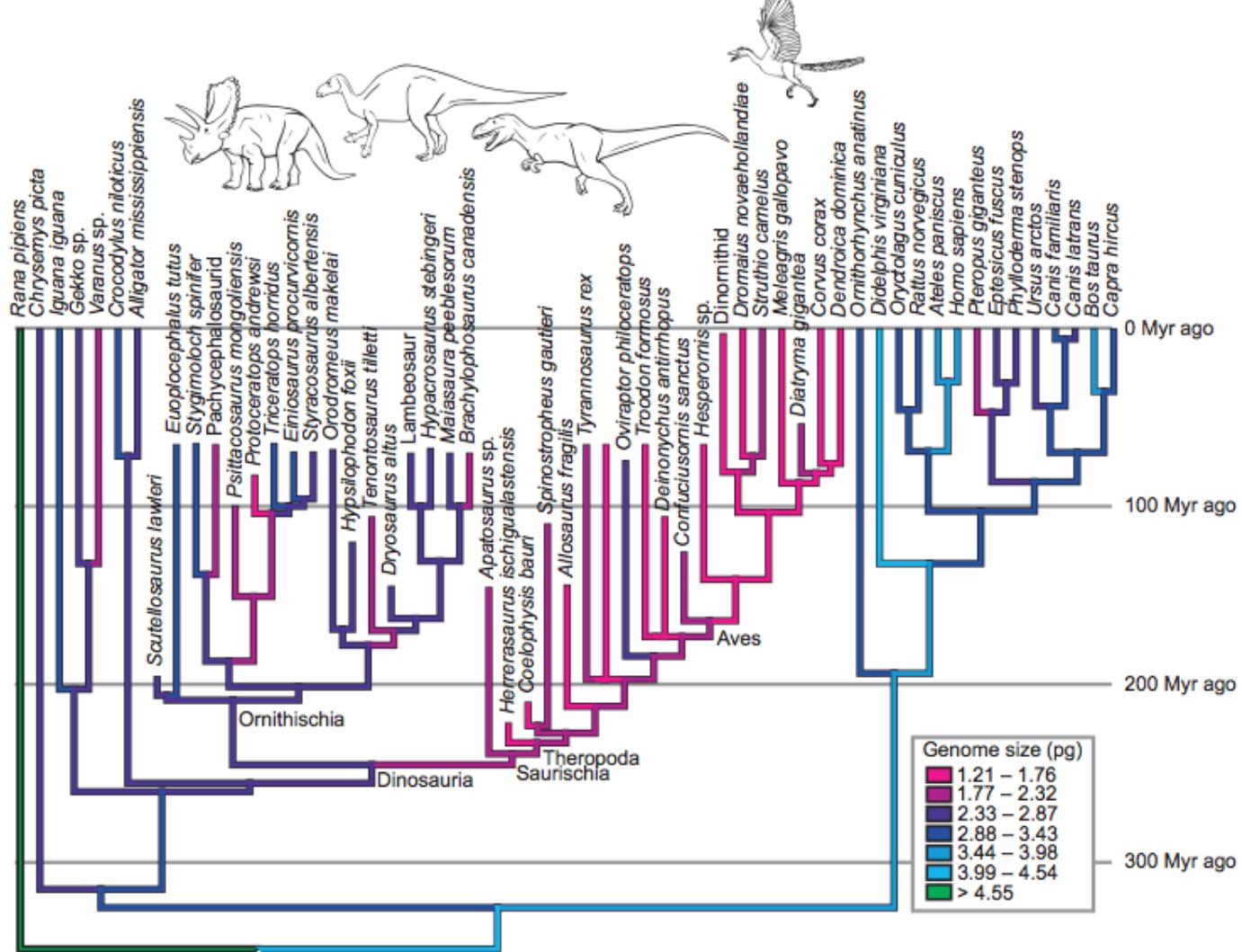
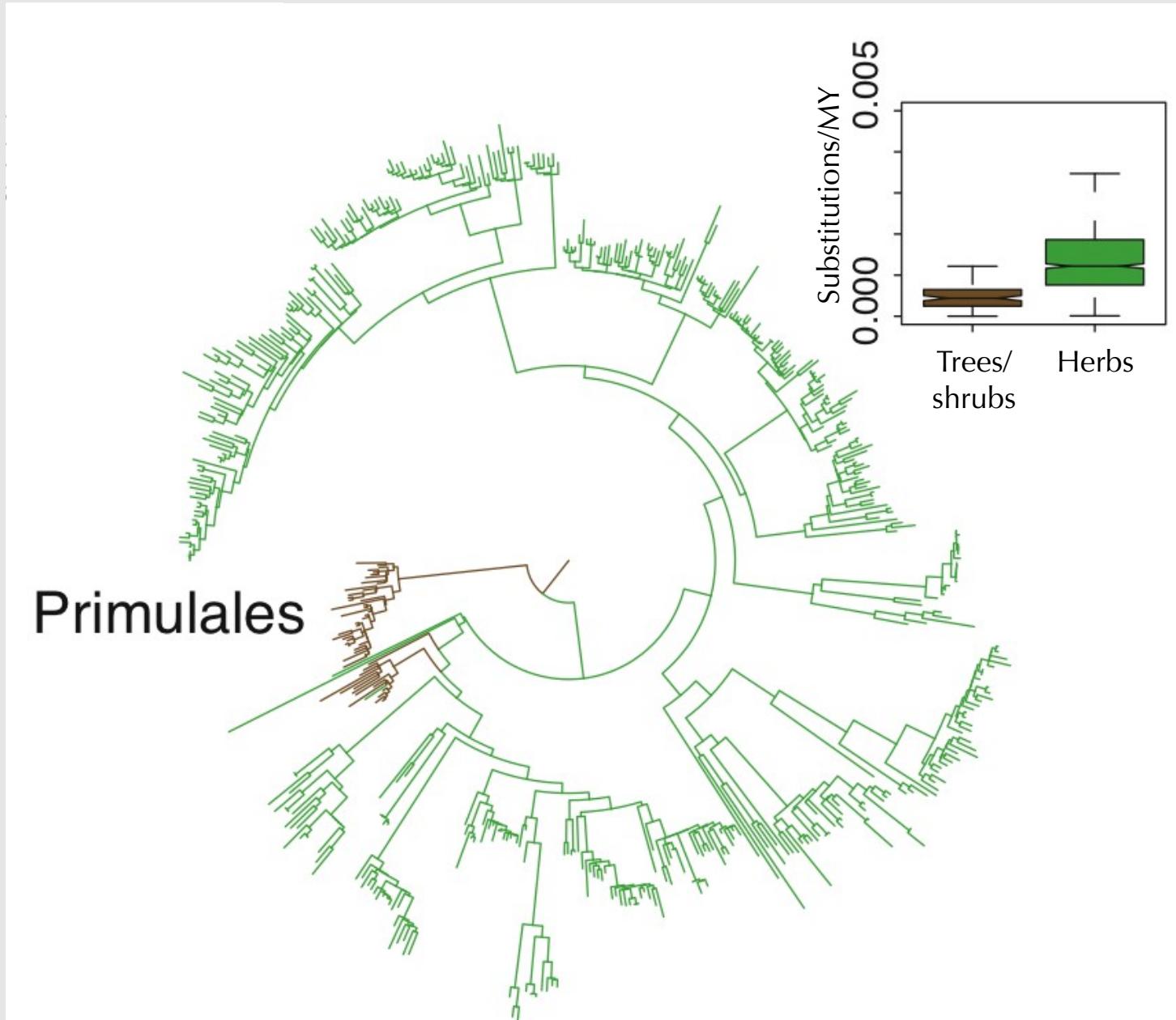


Figure 2 | Haplod genome size (mean of posterior predictive distribution) mapped onto a phylogeny shows a reduction within saurischian dinosaurs, the lineage to which birds belong. Myr, million years.

Here we use a novel bayesian comparative method to show that bone-cell size correlates well with genome size in extant vertebrates, and hence use this relationship to estimate the genome sizes of 31 species of extinct dinosaur, including several species of extinct birds. Our results indicate that the small genomes typically associated with avian flight evolved in the saurischian dinosaur lineage between 230 and 250 million years ago, long before this lineage gave rise to the first birds. By comparison,

ornithischian dinosaurs are inferred to have had much larger genomes, which were probably typical for ancestral Dinosauria. Using comparative genomic data, we estimate that genome-wide interspersed mobile elements, a class of repetitive DNA, comprised 5-12% of the total genome size in the saurischian dinosaur lineage, but was 7-19% of total genome size in ornithischian dinosaurs, suggesting that repetitive elements became less active in the saurischian lineage.



Smith and Donoghue. Rates of Molecular Evolution Are Linked to Life History in Flowering Plants. *Science* (2008)

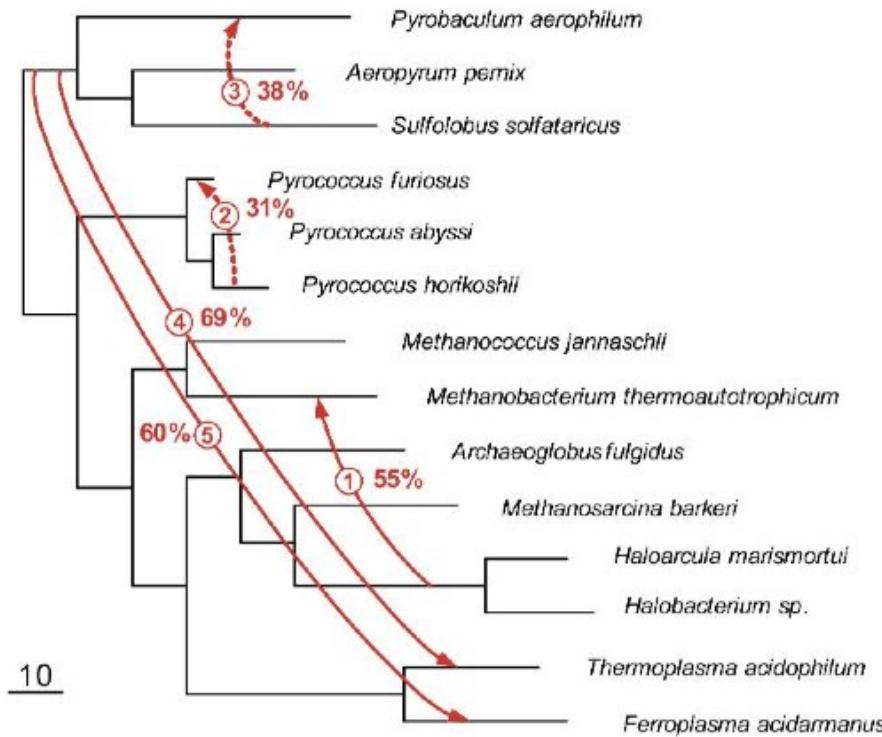
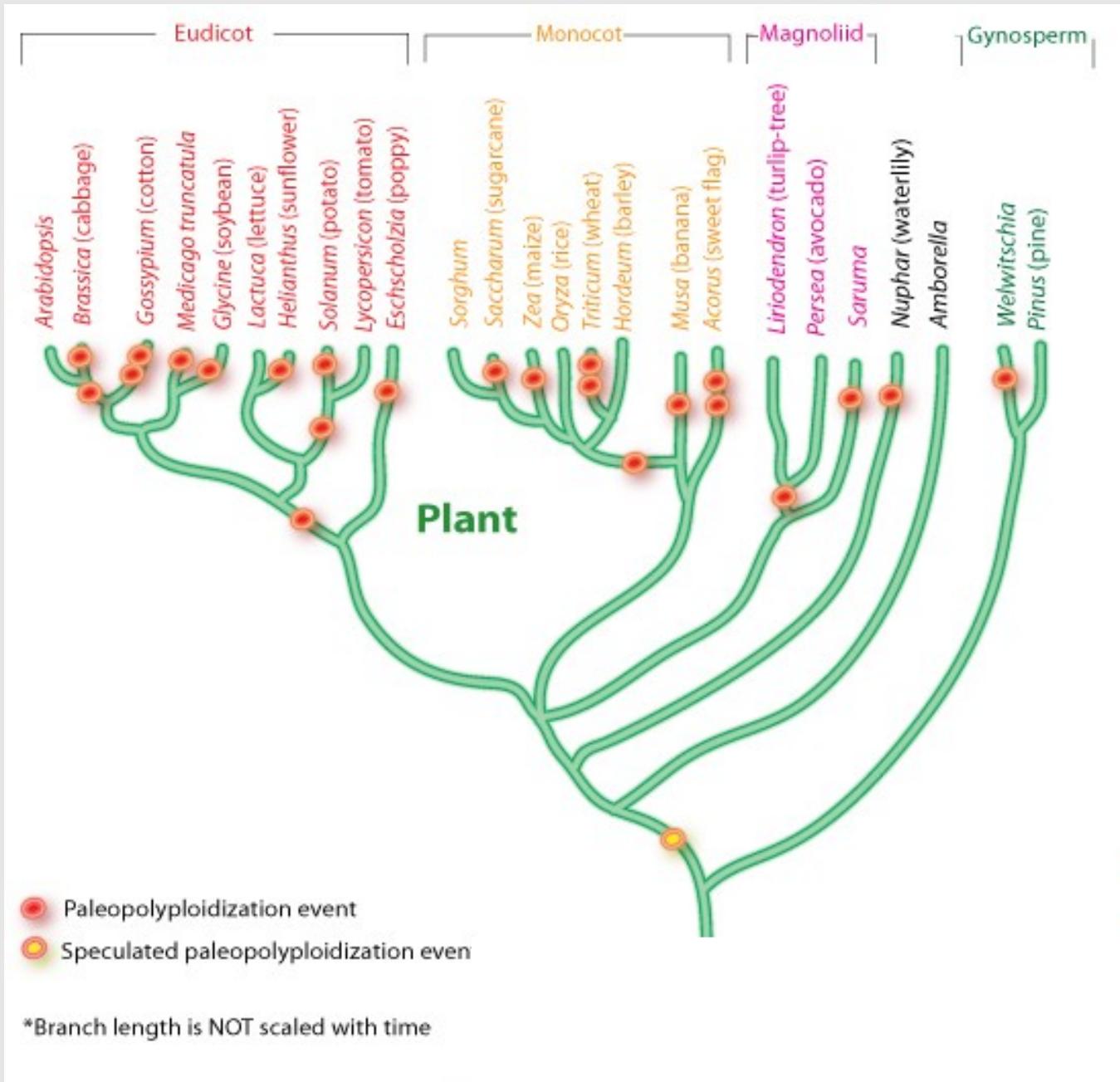


FIGURE 6. Species tree (Matte-Tailliez et al. 2002, fig. 1a) with 5 HGTs indicated by arrows. Numbers on HGTs indicate their order of inference. HGT bootstrap scores are indicated near to the numbers of the corresponding HGTs. Arrows 4 and 5 depict the HGTs between the clades of Thermoplasmatales and Crenarchaeota originally predicted by Matte-Tailliez et al. (2002). HGTs with bootstrap scores of 50% or less are depicted by dashed arrows.



Peter Zhang after Adams & Wendel, 2005; Cui et. al, 2006; Wolfe, 2001.

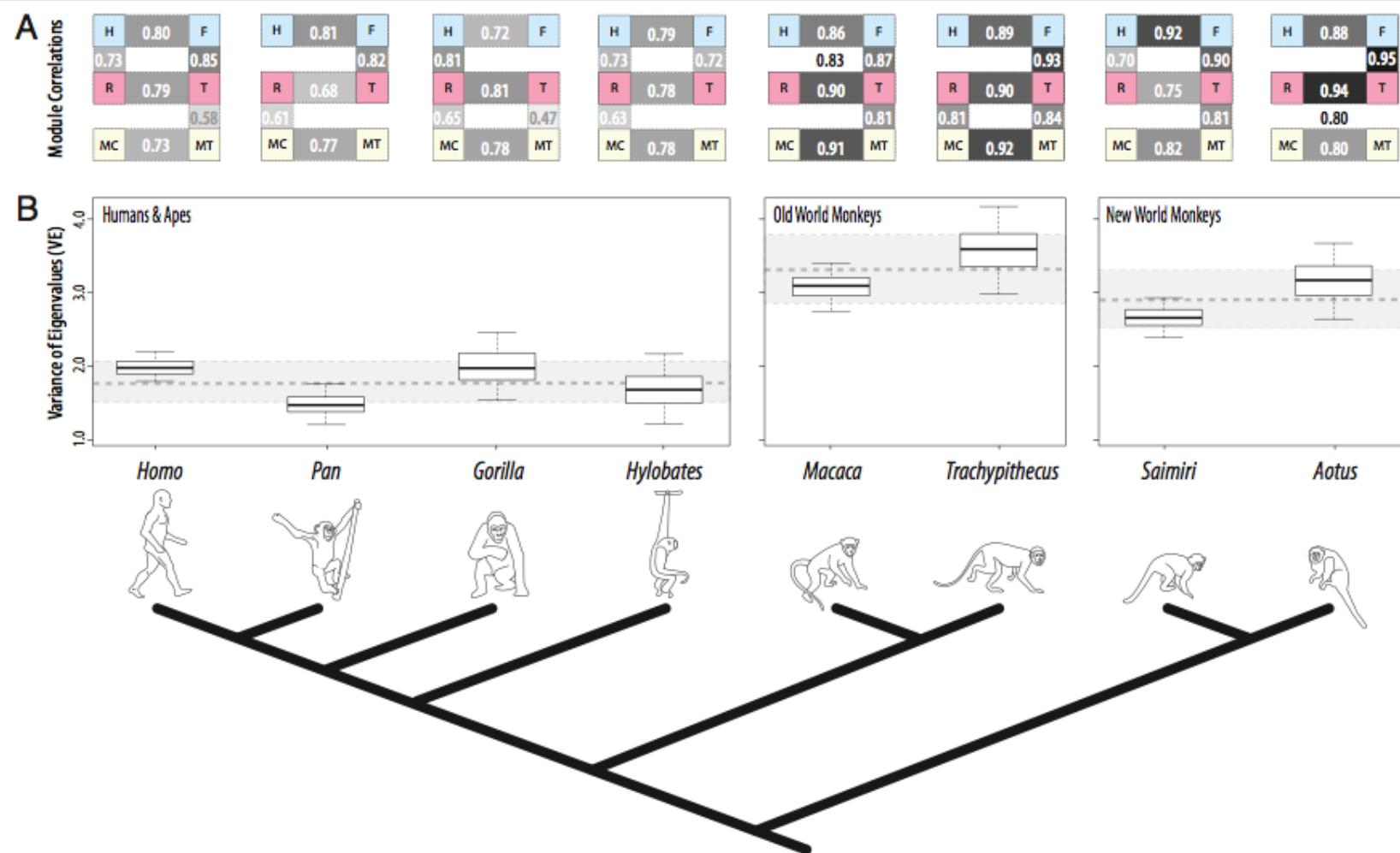


Fig. 2. (A) Anthropoid primates exhibit consistent developmental modularity of limbs. Within-limb modularity reflects varying functional signals (e.g., bipedalism in the hindlimb of humans and forelimb suspension in apes). Modules (partial correlations $P < 0.05$) are illustrated as boxes between elements. Modules are shaded relative to the strength of the estimated Fisher-z transformed correlation. Estimated Pearson correlation coefficient is shown. (MC, metacarpal; MT, metatarsal; R, radius; T, tibia; H, humerus; F, femur). Species arranged by phylogenetic relationship as shown at Bottom. (B) Humans are significantly less integrated compared to quadrupedal monkeys and similar to apes, indicating that reductions to integration and more independently evolvable limbs characterize both fossil hominins and hominoids. Box plots show the lower and upper quartile, median of resampled eigenvalue variance (VE) (10,000 replicates). Whiskers indicate the 95% confidence limit of the estimate. Dashed lines and shaded boxes show the average VE for hominoids ($=1.79$), cercopithecoids ($=3.32$), and ceboids ($=2.90$) and the 95% confidence interval, respectively.

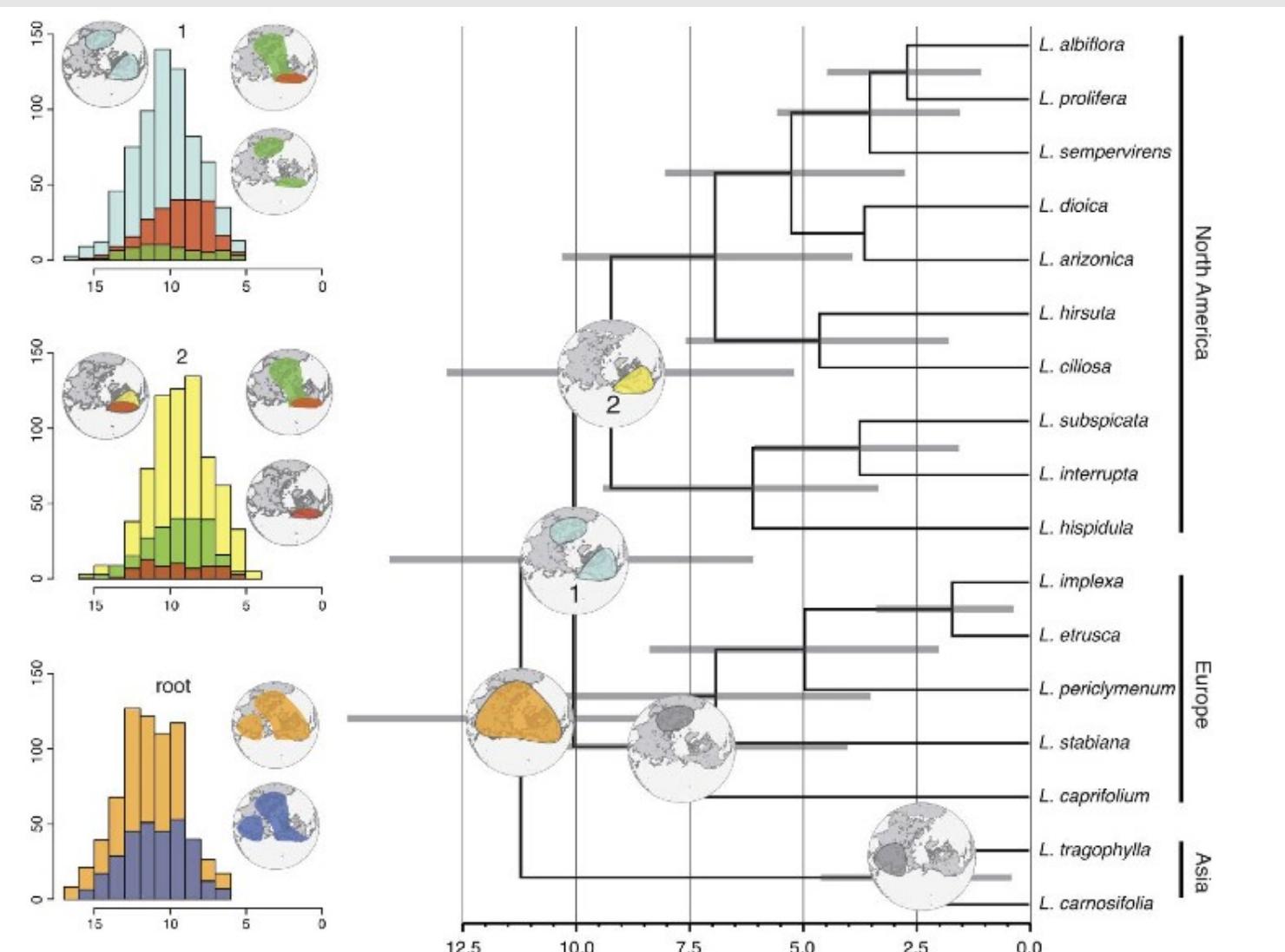
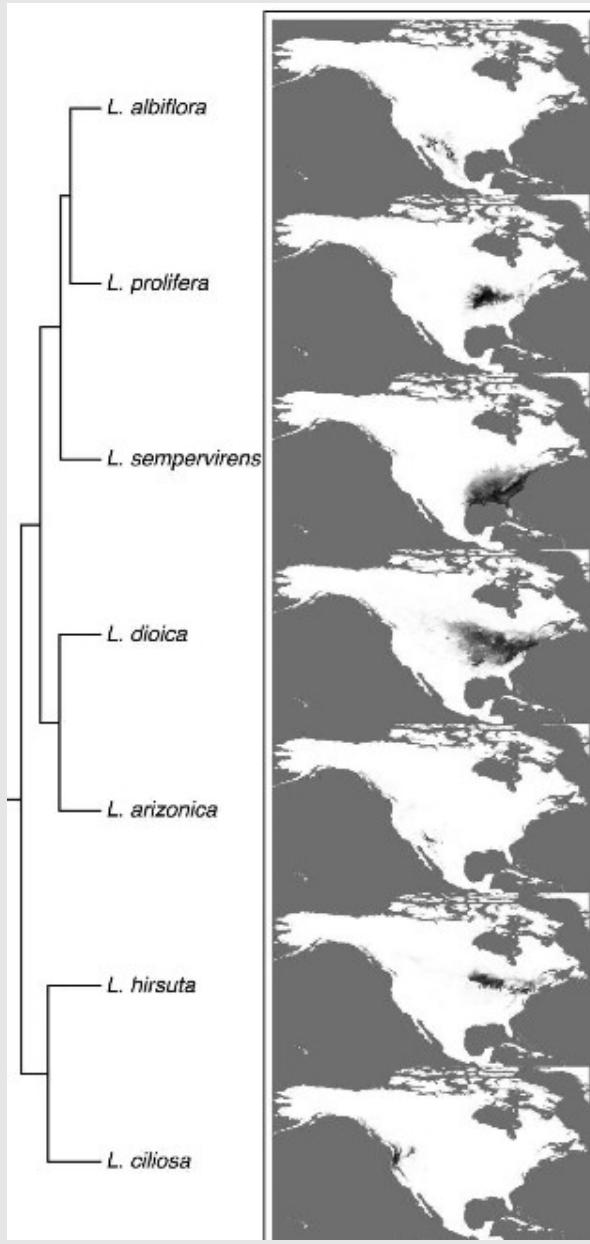


FIGURE 2. Historical biogeography of the *Caprifolium* clade of *Lonicera*. The phylogeny is the maximum credibility clade from the Bayesian divergence time analysis with *L. flava* and *L. pilosa* included. Inferred ancestral ranges for 5 major well-supported clades are shown. Speciation scenarios are shown in the histograms and associated maps. The x-axis on the phylogeny and in the histograms represents time in millions of years; histograms display variance in the inferred ancestral area as a function of time.



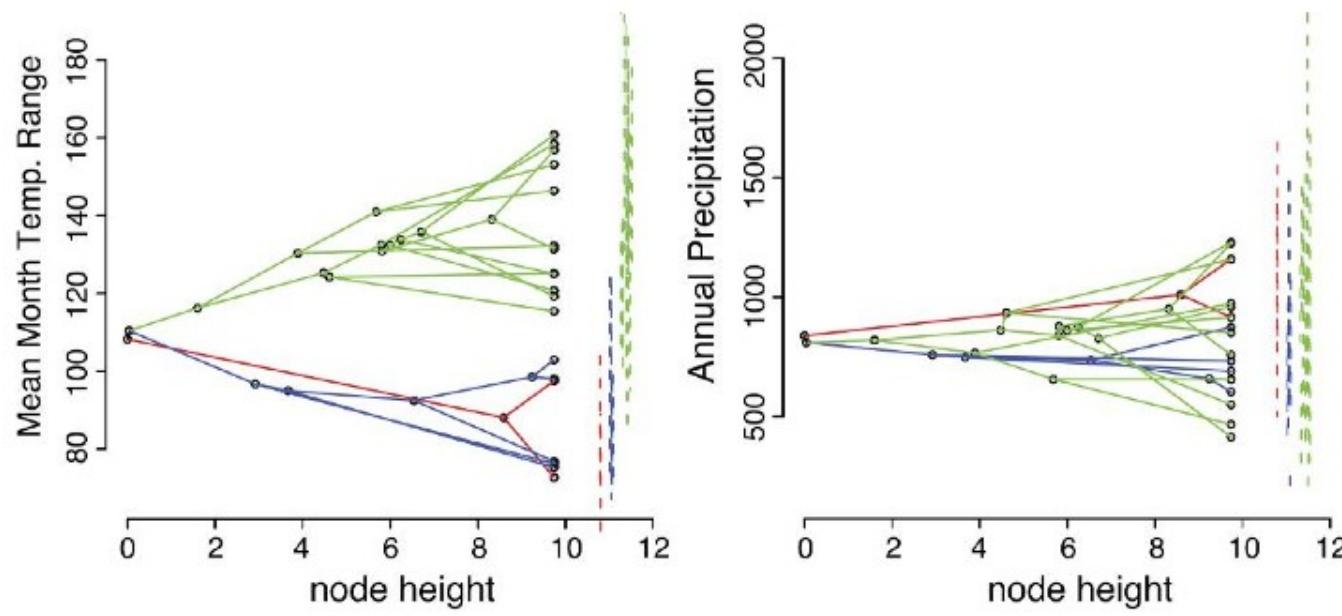
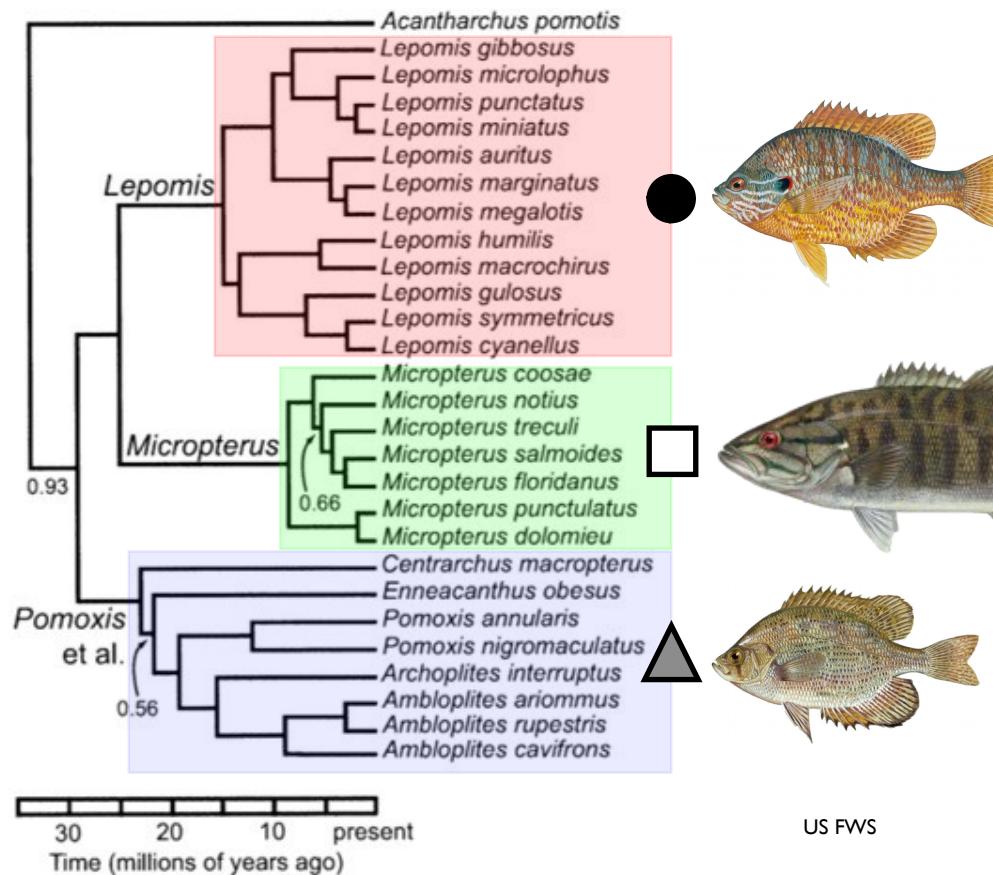
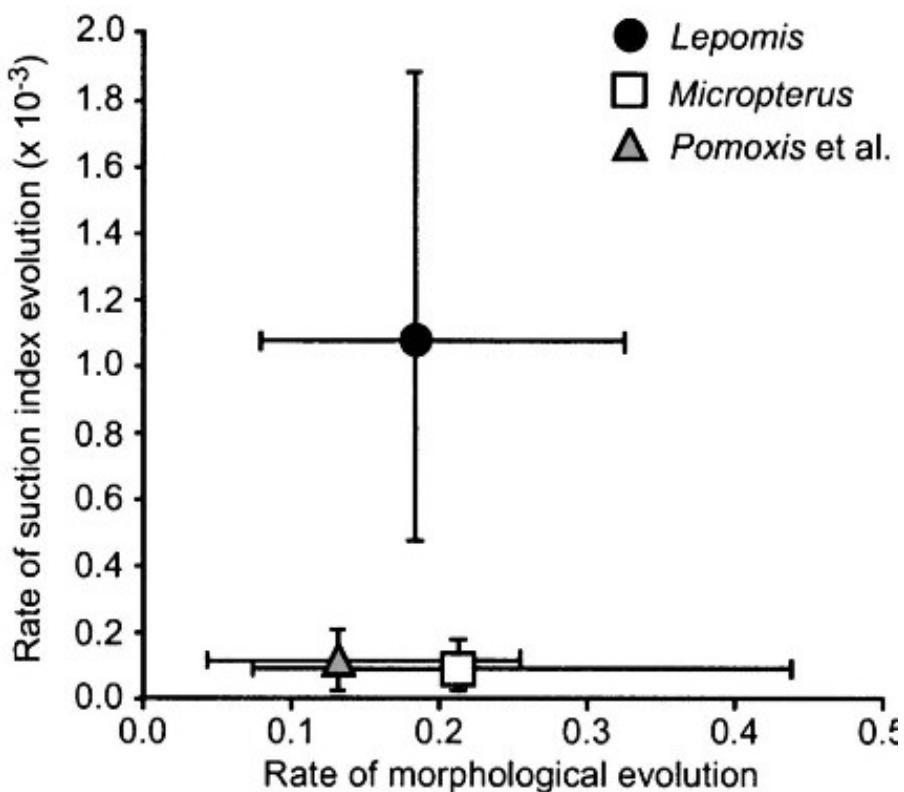


FIGURE 5. Examples of ancestral reconstructions of climate niches. Green branches represent species in North America, red branches in Asia, and blue branches in Europe. Extant disparity is represented on the right; the x-axis represents time. The plot on the left represents mean monthly temperature range; this exhibits greater differences among clades than within clades, and greater niche conservatism. The plot on the right represents annual precipitation; this exhibits greater differences within clades than between clades, and less niche conservatism. Results for all climate traits are presented in the Supplemental Figures.

DISCORDANCE BETWEEN MORPHOLOGICAL AND MECHANICAL DIVERSITY IN THE FEEDING MECHANISM OF CENTRARCHID FISHES

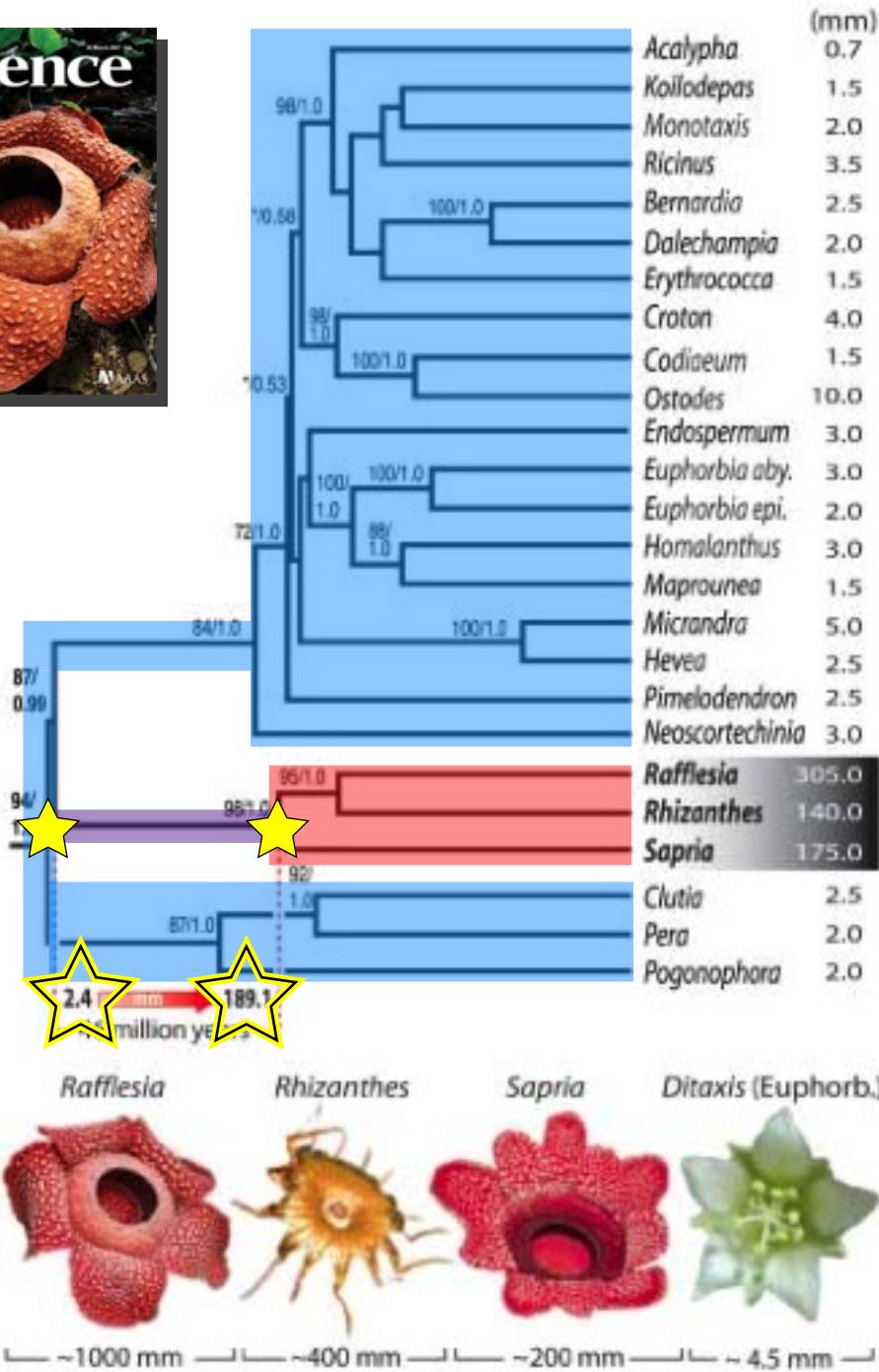
DAVID C. COLLAR^{1,2} AND PETER C. WAINWRIGHT^{1,3}



US FWS



Science



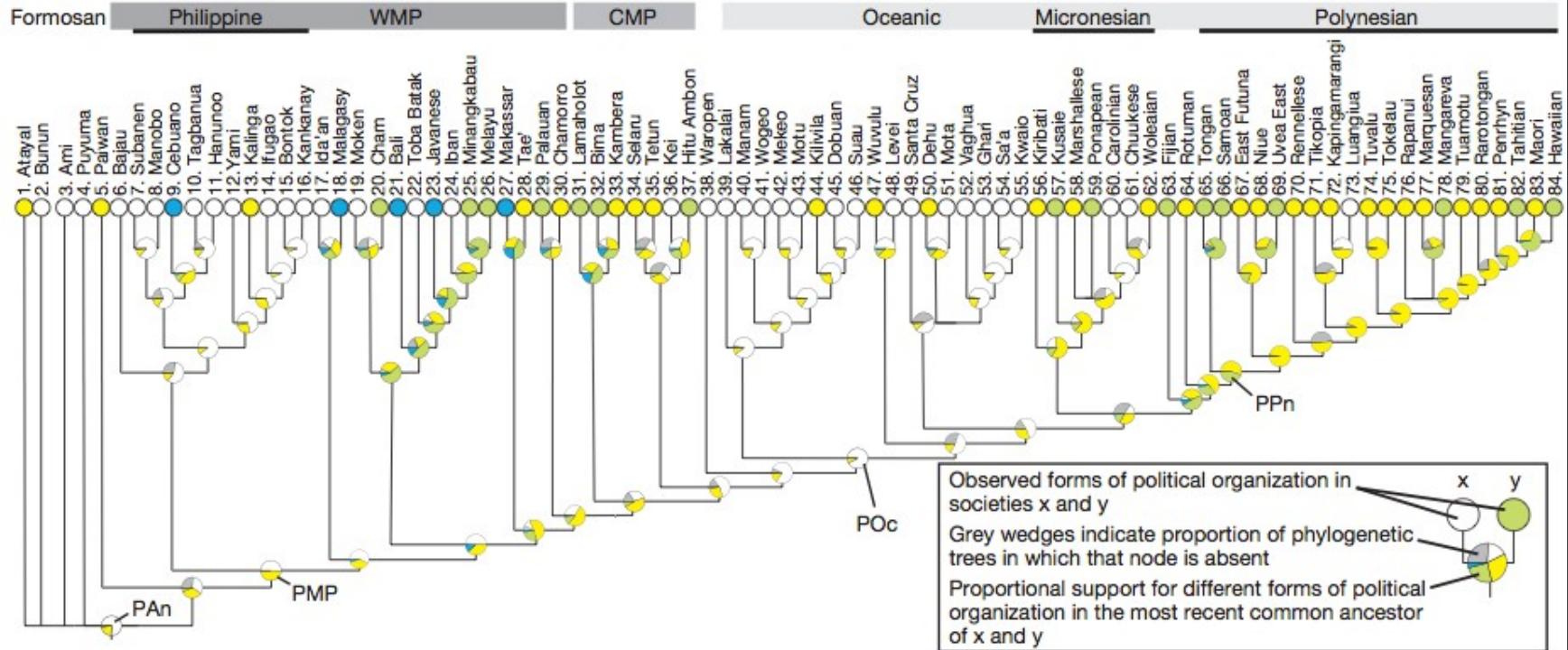
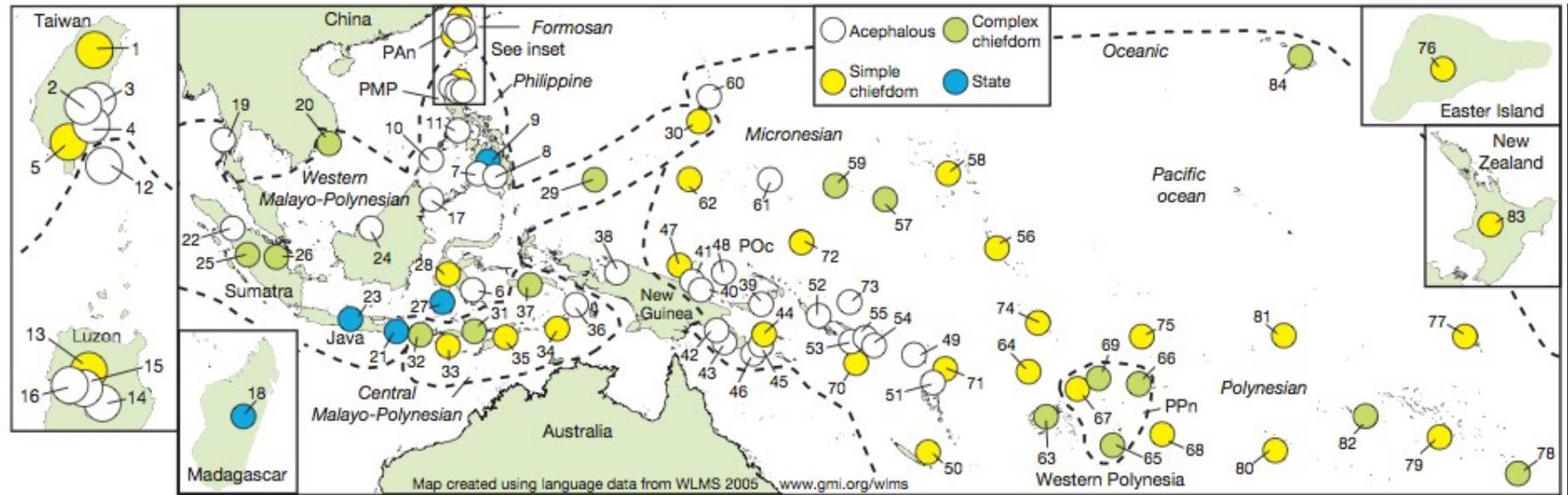
Floral Gigantism in Rafflesiaceae

Charles C. Davis,^{1*} Maribeth Latvis,¹ Daniel L. Nickrent,² Kenneth J. Wurdack,³ David A. Baum⁴

◆ Estimate rates
A, B, and C

◆ Stretch tree
based on these
rates

◆ Estimate states
on stretched tree



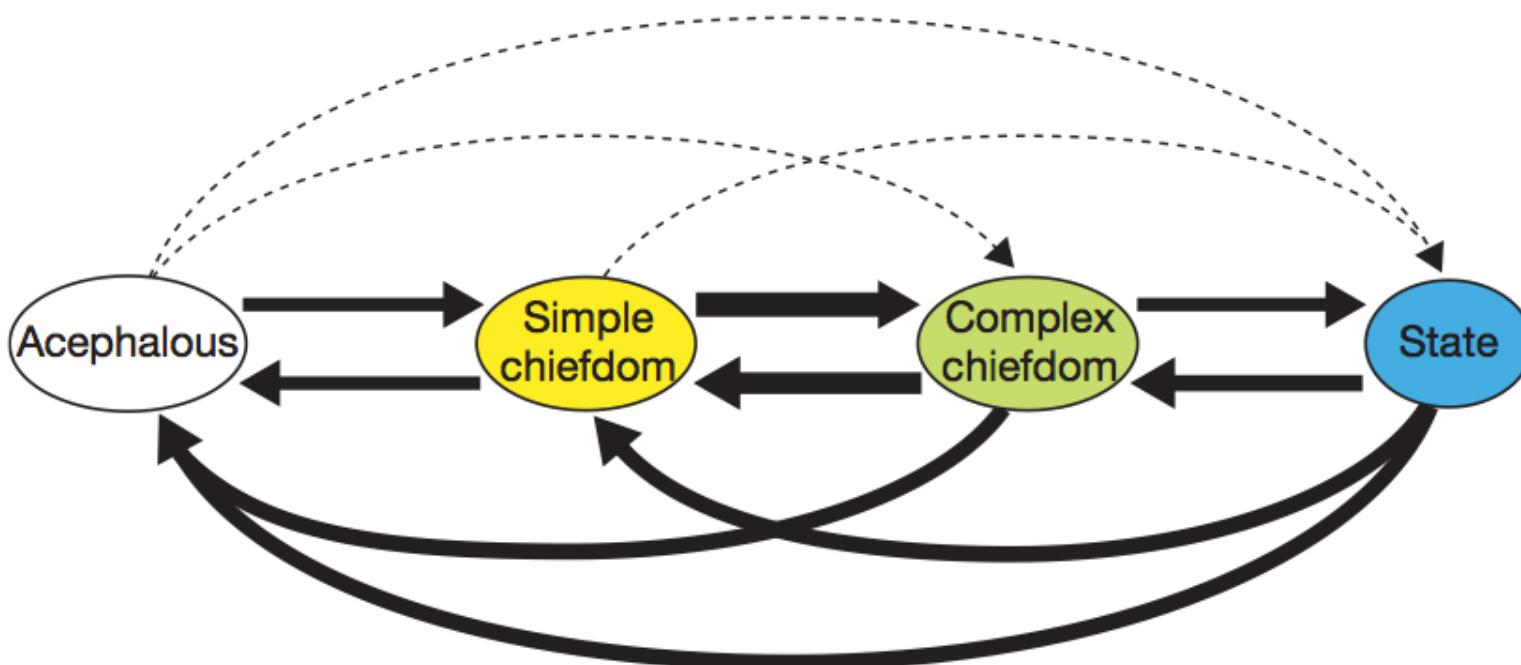
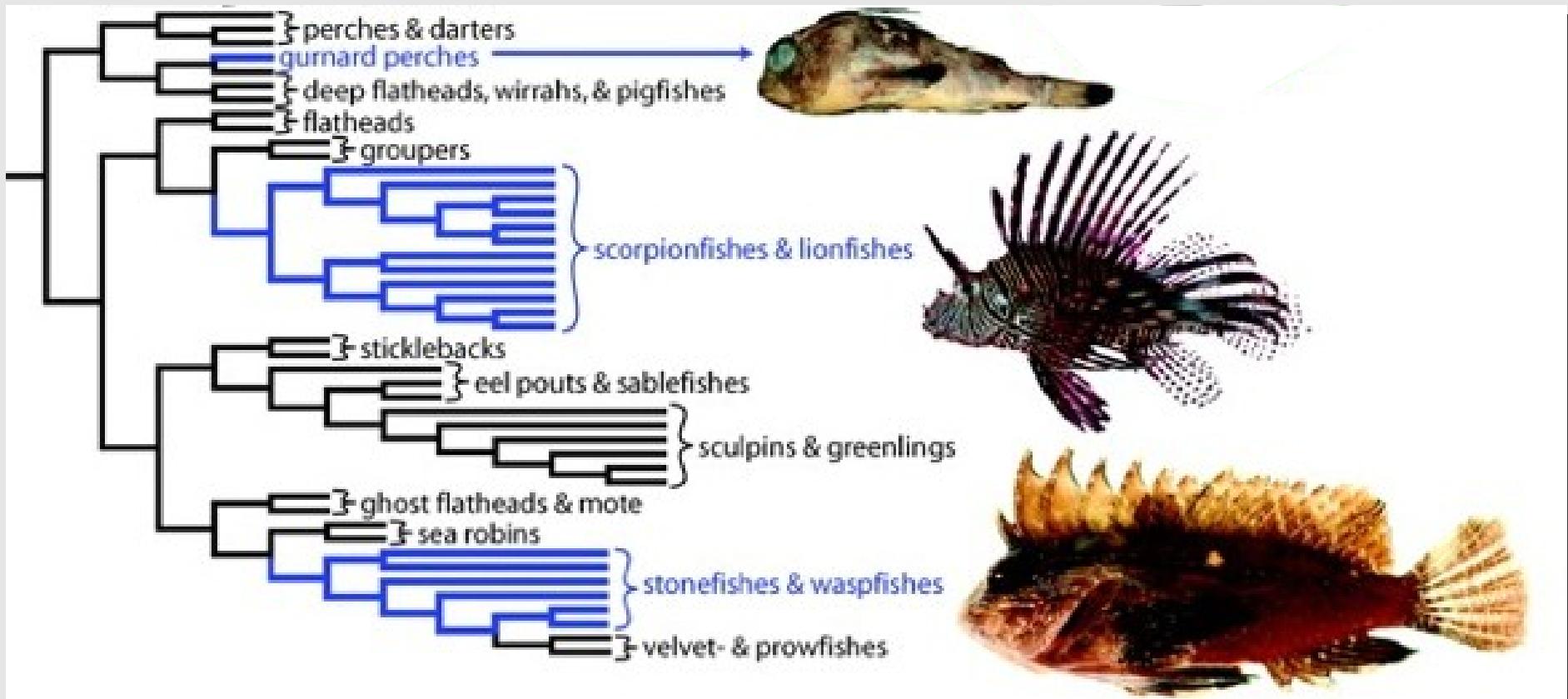


Figure 3 | Estimated rate parameters for transitions between different forms of political organization under the RJMCMC analysis. Arrow thickness is proportional to the mean estimated value of the associated rate parameter. Transitions from acephalous society to state, acephalous society to complex chiefdom and simple chiefdom to state are represented by a dotted line as they are often estimated as zero (Supplementary Methods, Supplementary Table 4 and Supplementary Fig. 3). These results are consistent with the idea that political evolution has followed a sequence of incremental increases in hierarchical organization.



Smith &
Wheeler 2006

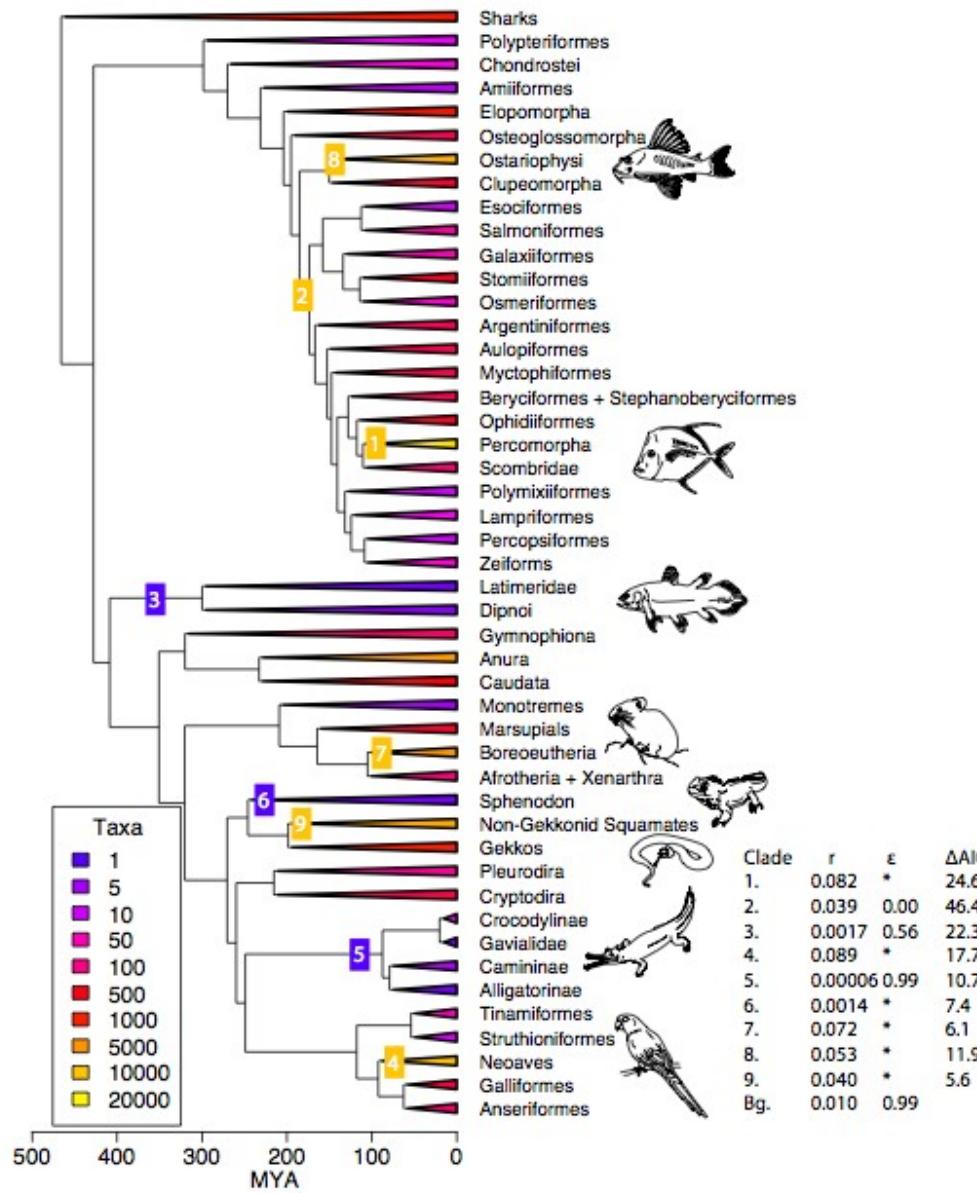


Fig. 1. Diversity tree for analyses of lineage diversification in vertebrates. Clades are collapsed to 47 representative stem lineages and colored by extant species diversity. Clades with unusual diversification rates are denoted with numbers that indicate the order in which rate shifts were added by the stepwise AIC procedure; yellow and blue squares denote diverse and impoverished clades, respectively, compared with background rates. Estimates for net diversification rate ($r = \lambda - \mu$) and relative extinction rate ($\epsilon = \mu/\lambda$) are included in the lower right table. Relative extinction can be calculated only when at least part of the subclade is resolved [see Rabosky et al. (13)]. Asterisks indicate subclades where ϵ values could not be estimated for this reason.

What questions could
we address using these
techniques?