

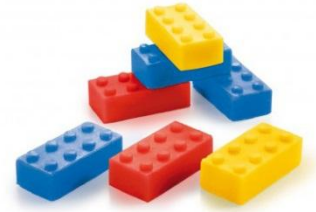
# **WORKSHOP IN ADVANCED BAYESIAN PHYLOGENETICS**

**NOV 17-21, 2014, ADELAIDE**



**WHY CALIBRATING?**  
**FIXED-RATE CALIBRATION**  
**NODE CALIBRATION**  
**TIP DATING**

# WHY CALIBRATING?



**data (DNA, protein etc.)**



**substitution length**

**clock model**

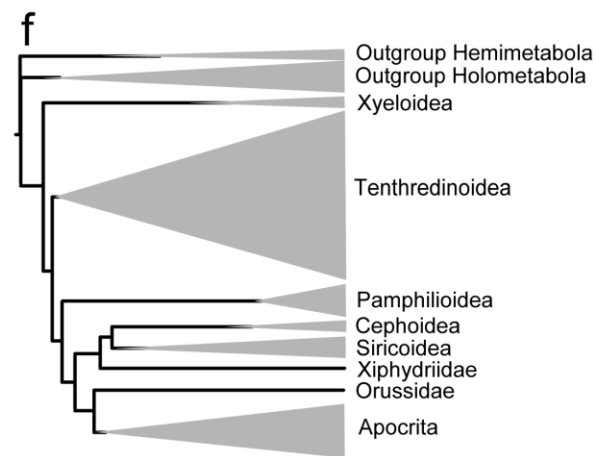
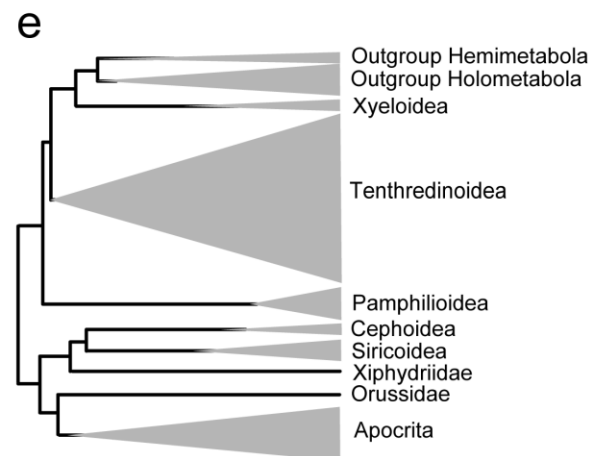
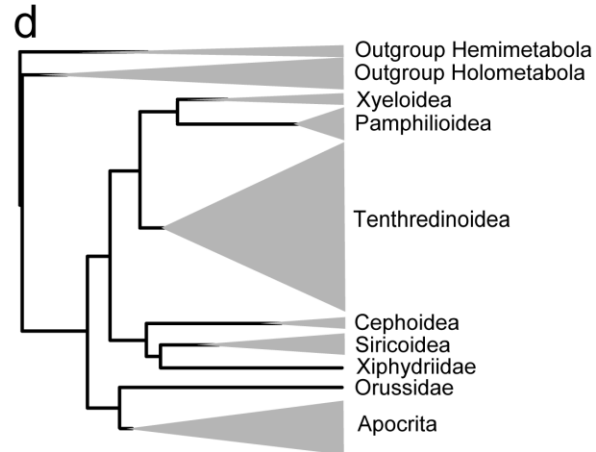
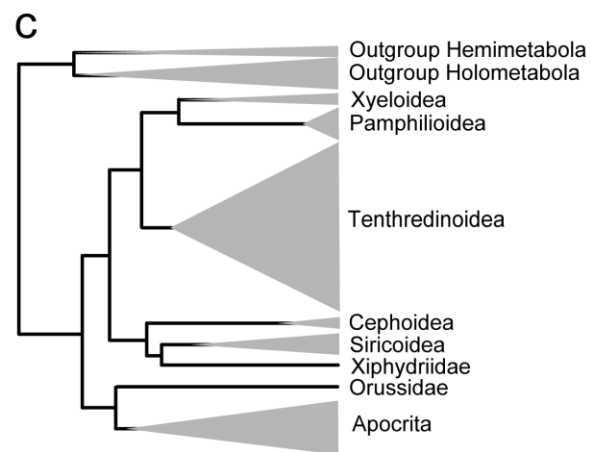
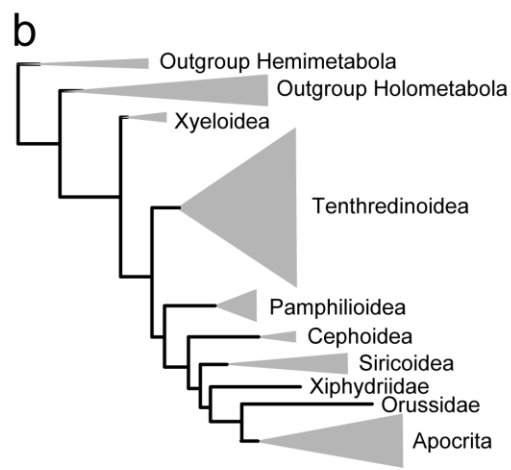
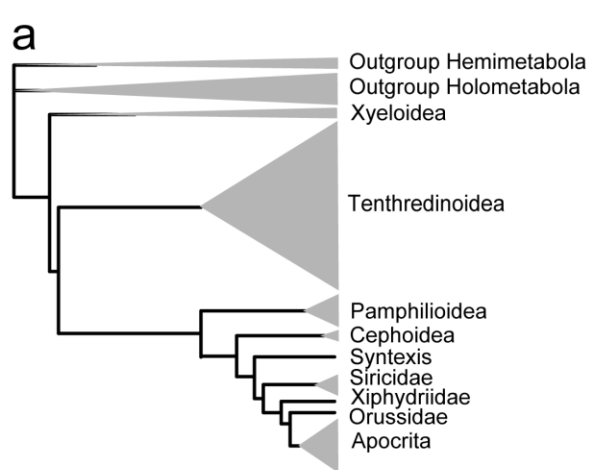


**time length (in subst.)**

**calibration**



**time length (in years o.s.)**



**a:** morphology

**b:** non-clock

**c:** strict clock

**d:** strict clock,  
constraint on  
Holometabola

**e:** relaxed clock

**f:** relaxed clock +  
rooting constraint

# WHY CALIBRATING?



what do branches mean in clock trees?

time length<sub>branch i</sub> =

substitution length<sub>branch i</sub> / rate<sub>branch i</sub>

# CALIBRATION

## several sources of information about time scale

- molecular clock **rate estimates** from other study
- **node calibration**:
  - fossils
  - **biogeography**, e.g. continental breakups
  - secondary calibration: previous dating studies
- **tip dating** (viruses, ancient DNA, total-evidence)

# FIXED-RATE CALIBRATION

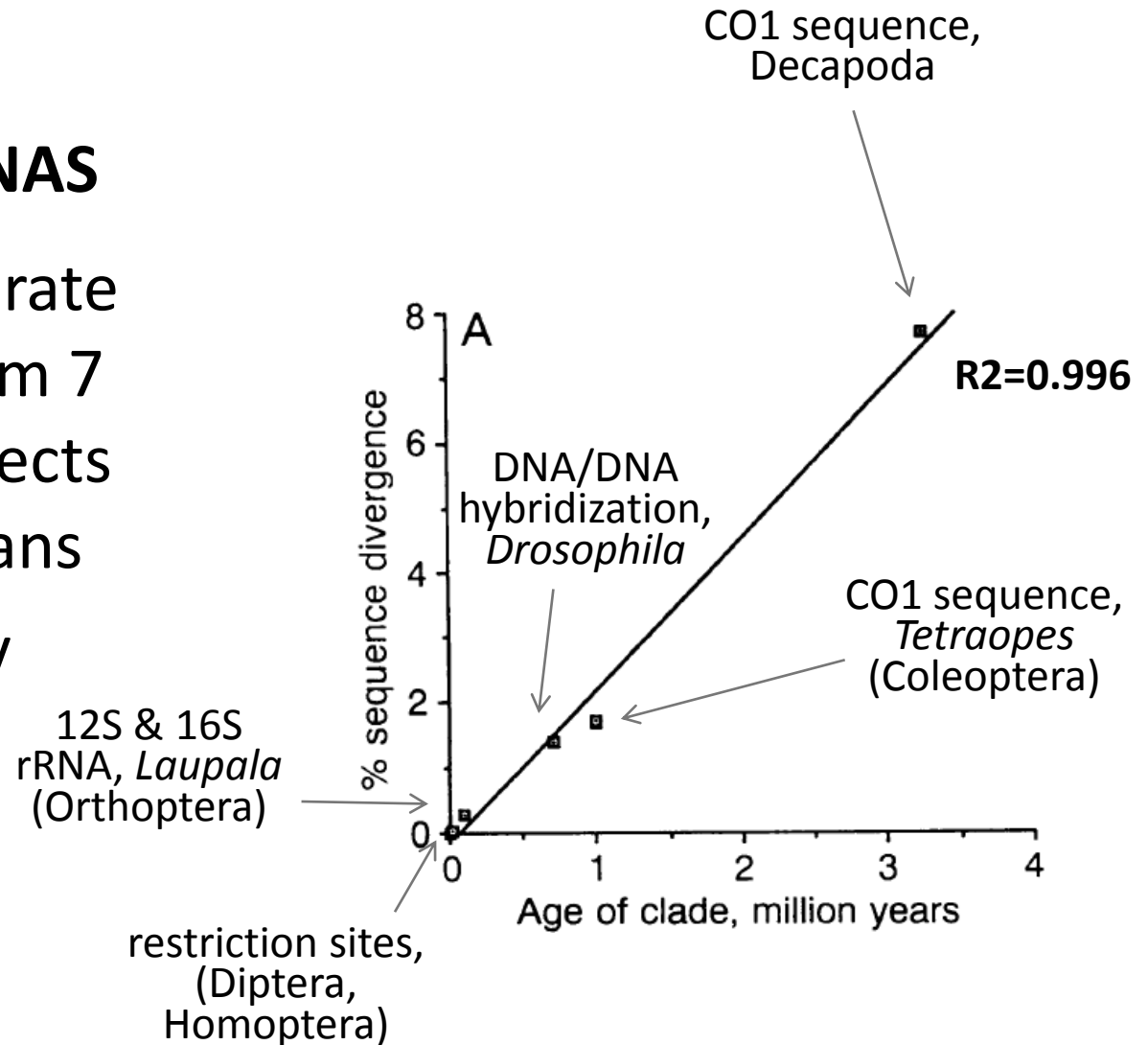
**molecular clock** **rate estimates** from other study

- e.g., in insects, mtDNA evolves at a rate of  $\sim 2.3$  substitutions per million years (?)
- pretty inexact way: large variation in rates among lineages observed!
- only works if genes not saturated  $\rightarrow$  more recent time scales

# FIXED-RATE CALIBRATION

## Brower 1994, PNAS

- independent rate estimates from 7 studies of insects and crustaceans
- 2.34% per my



# FIXED-RATE CALIBRATION

## critique of fixed-rate approach

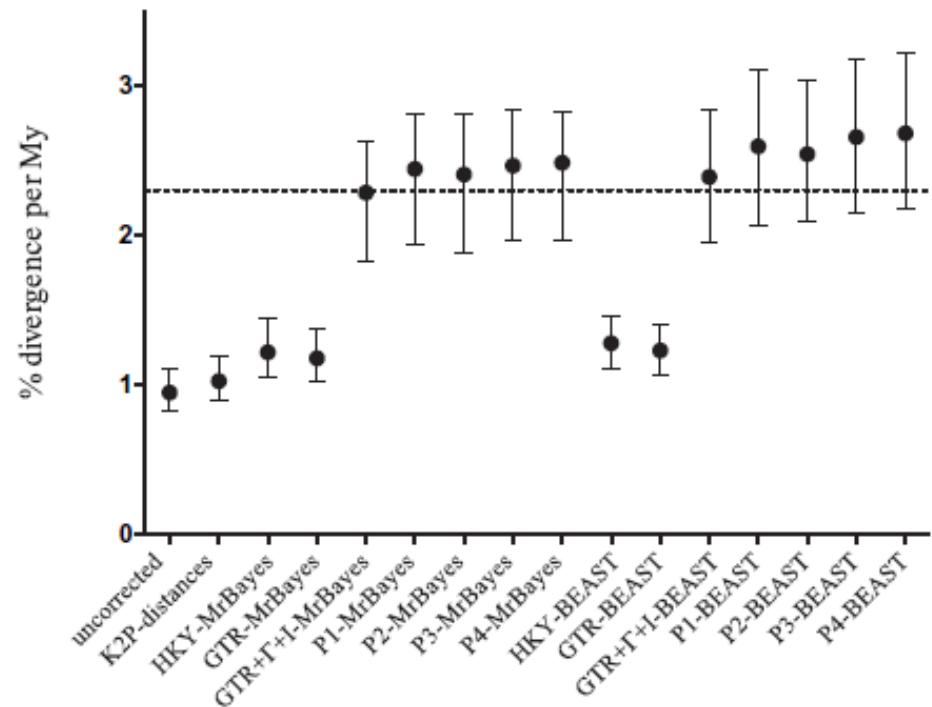
- lineage-specific effects
- gene-specific effects
- evolutionary model



# FIXED-RATE CALIBRATION

## critique of fixed-rate approach

- lineage-specific effects
- gene-specific effects
- evolutionary model

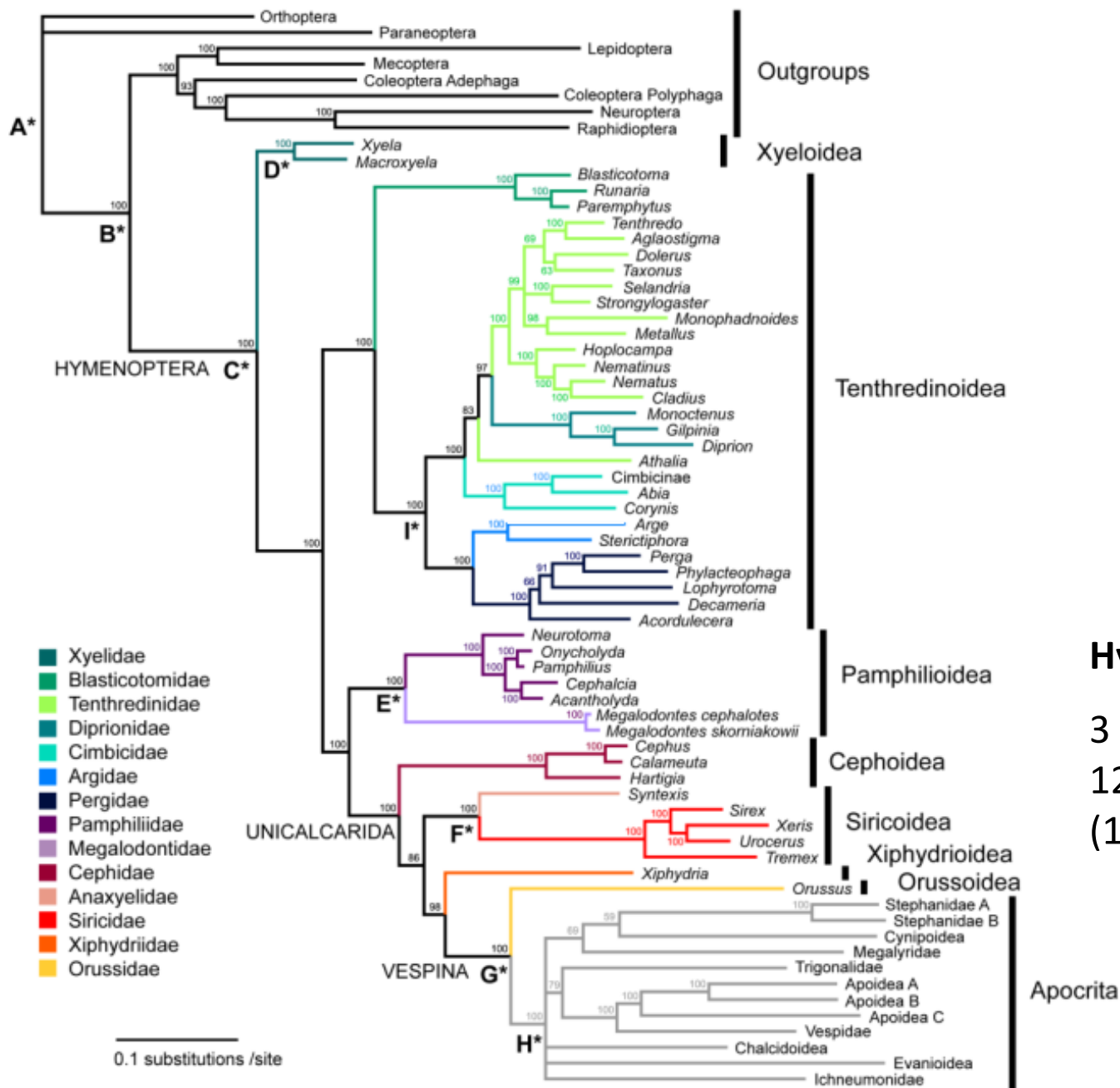


Papadopoulou et al. 2010, MBE

# FIXED-RATE CALIBRATION

## critique of fixed-rate approach

- lineage-specific effects
- gene-specific effects
- evolutionary model
- time-dependency: higher rates close to population level ( $< 1$  my) than above species level (delayed effect of purifying selection?)



## Hymenoptera

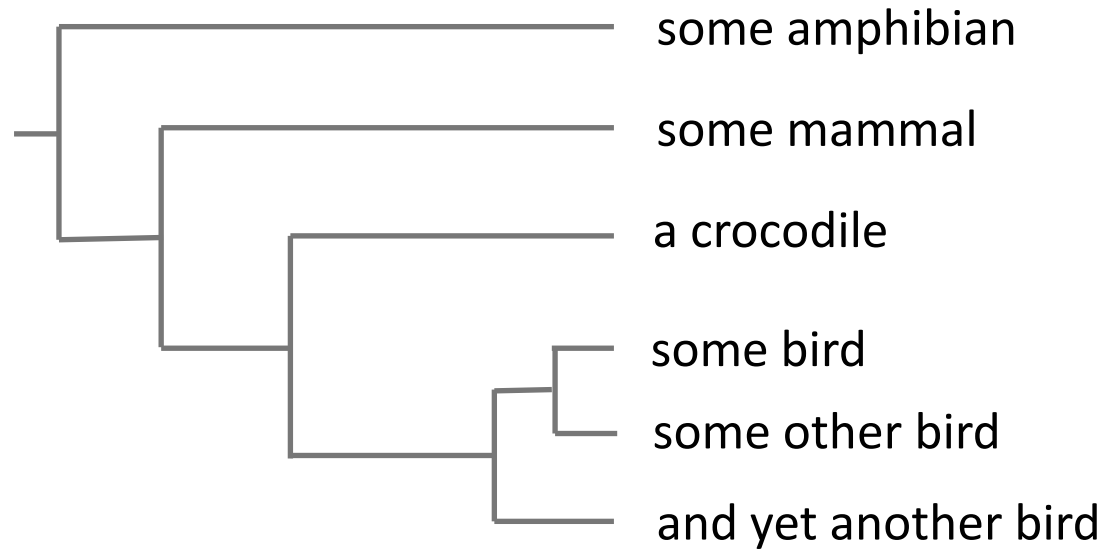
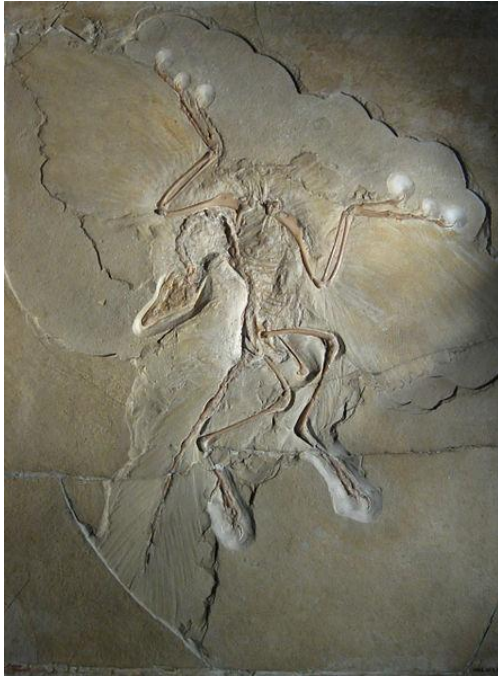
3 mitochondrial (CO1, 16S, 12S) and 4 nuclear genes (18S, 28S, EF1 $\alpha$  F1 & F2)

# NODE CALIBRATION

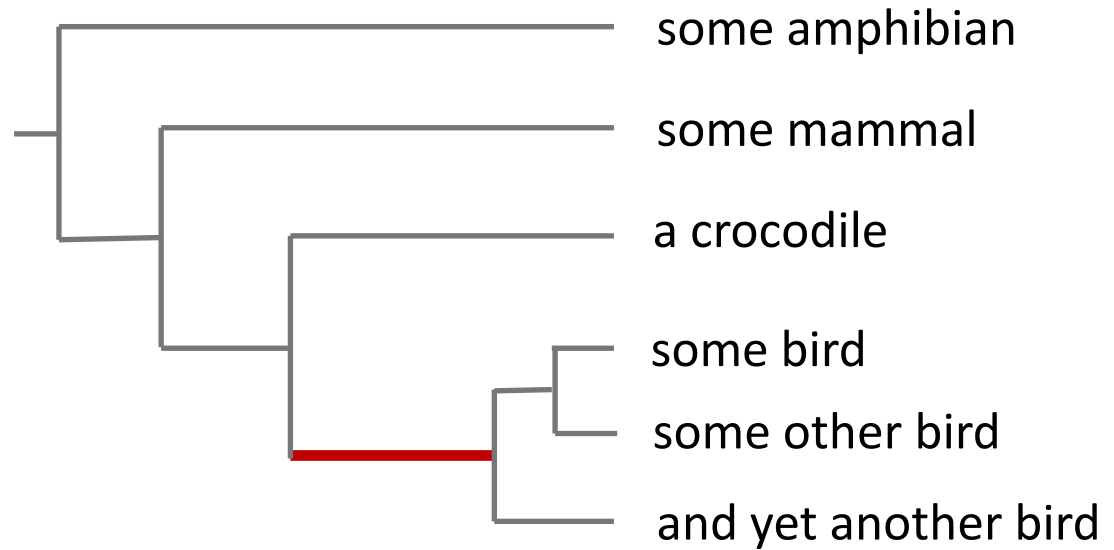
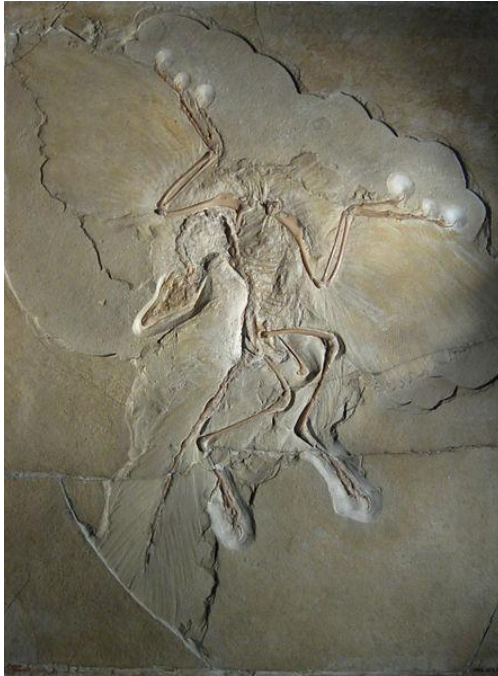
## fossils

- derive calibration point(s) from fossils

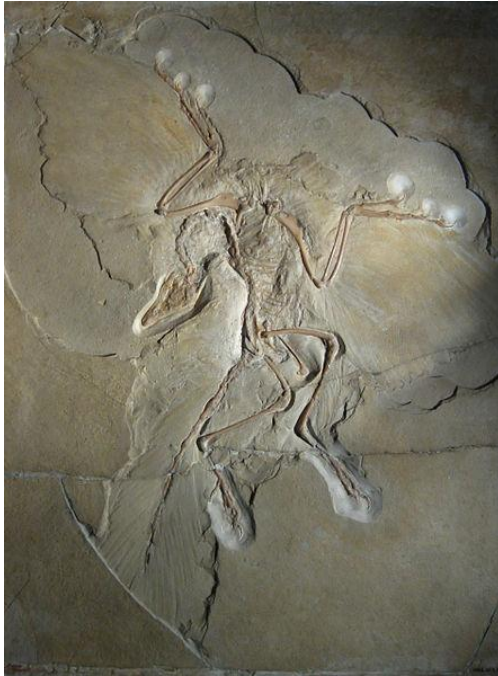
# FOSSIL CALIBRATION



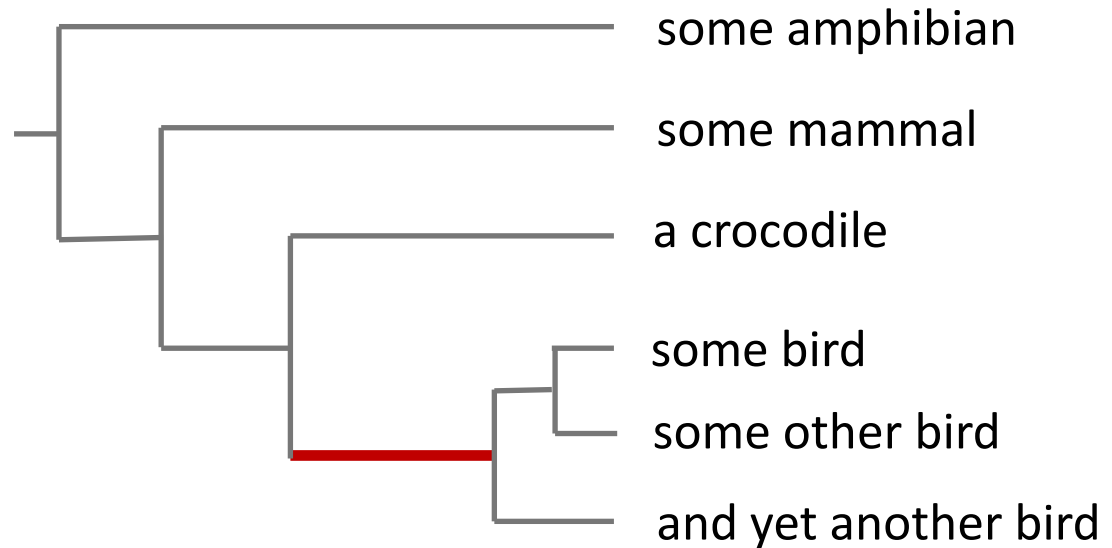
# FOSSIL CALIBRATION



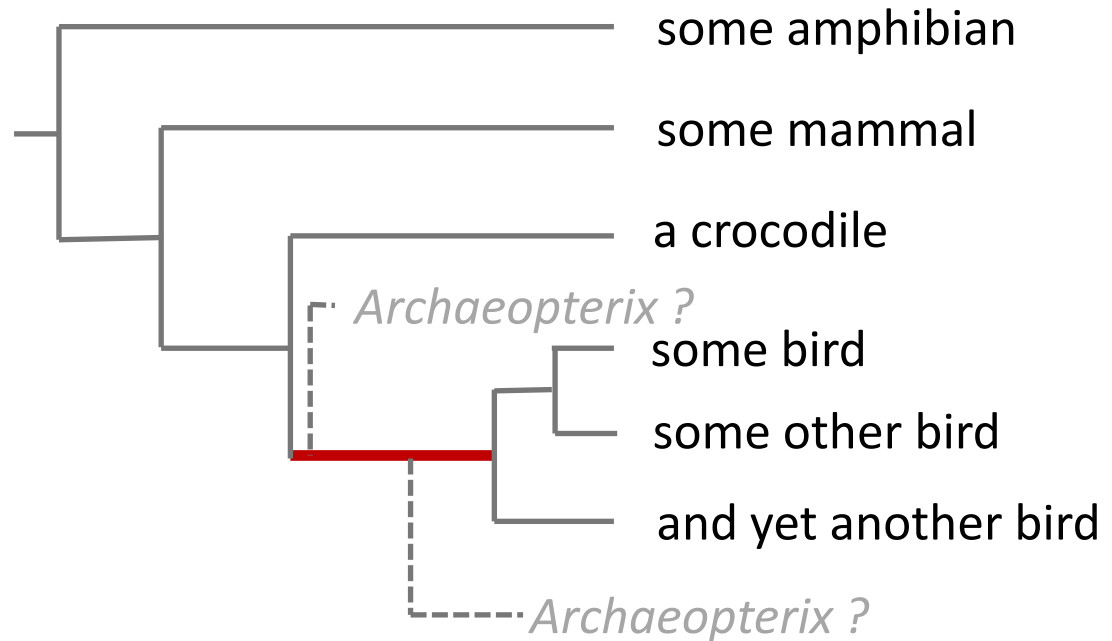
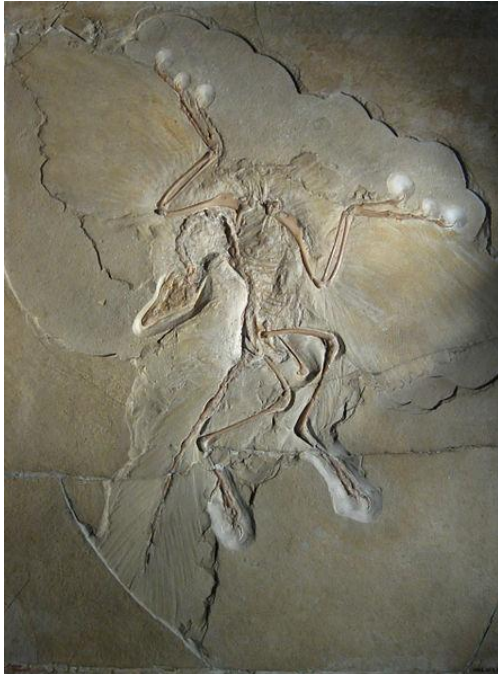
# FOSSIL CALIBRATION



minimum?  
or maximum?  
or both?

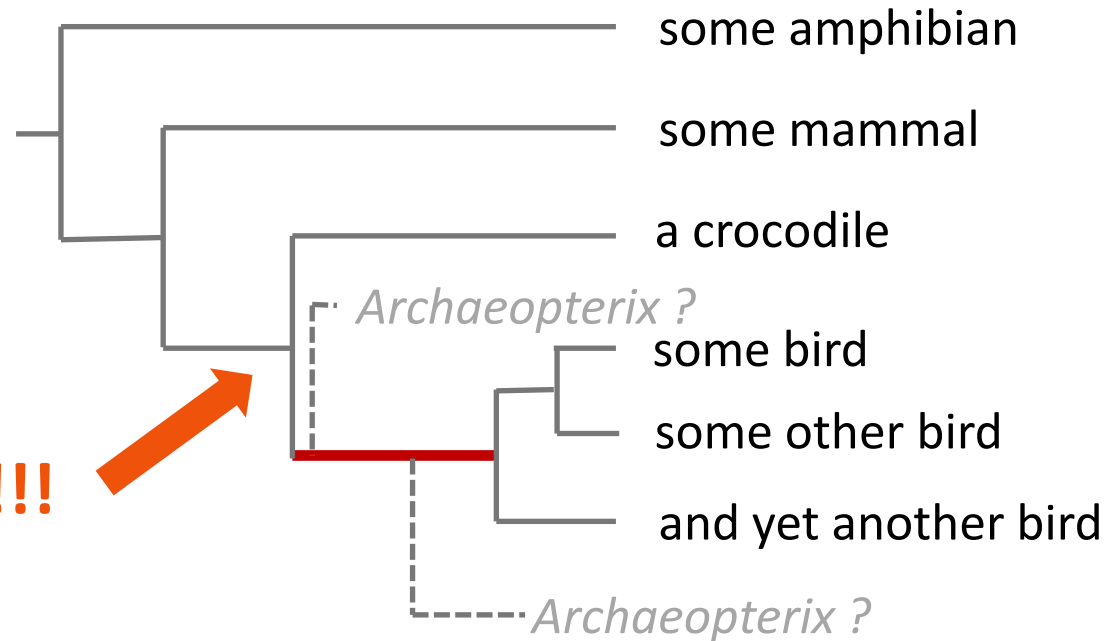
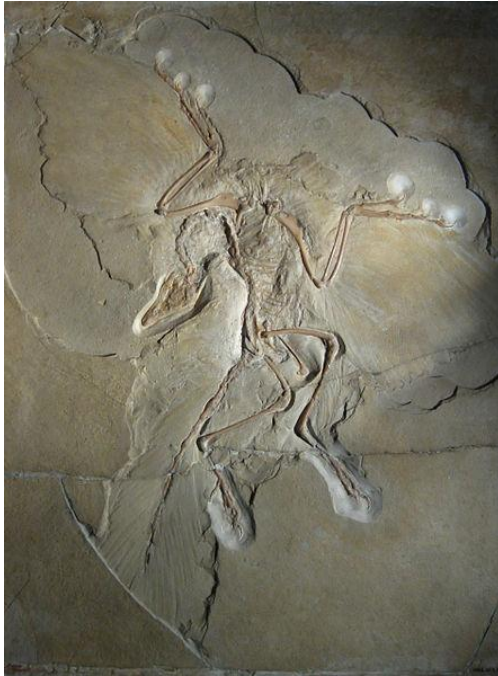


# FOSSIL CALIBRATION





# FOSSIL CALIBRATION



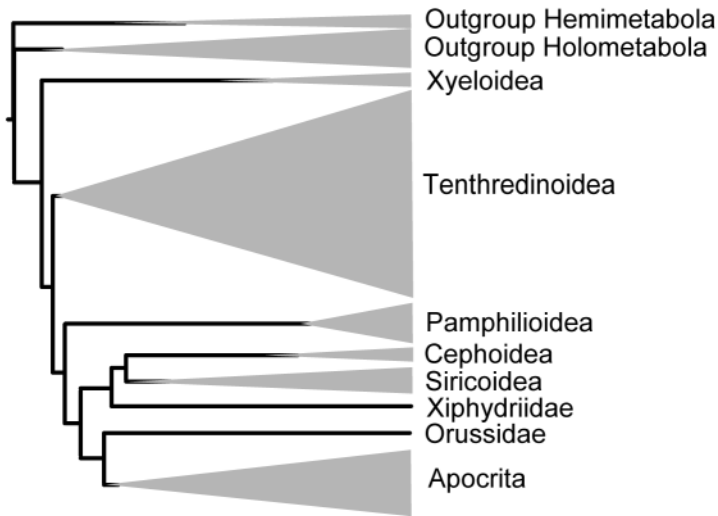
**minimum only!!!**

# FOSSIL CALIBRATION

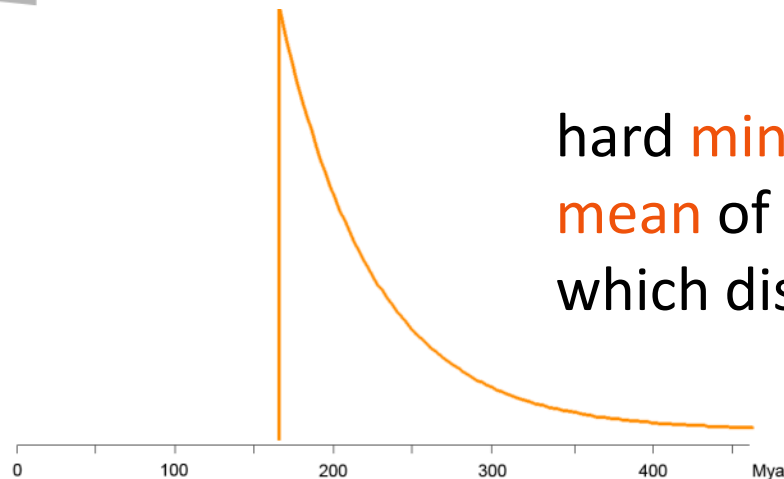
## fossils

- derive calibration point(s) from fossils
- give a **minimum** age only, and of the node **below** the group to which it belongs
- hard minimum, + usually an arbitrary probability distribution...

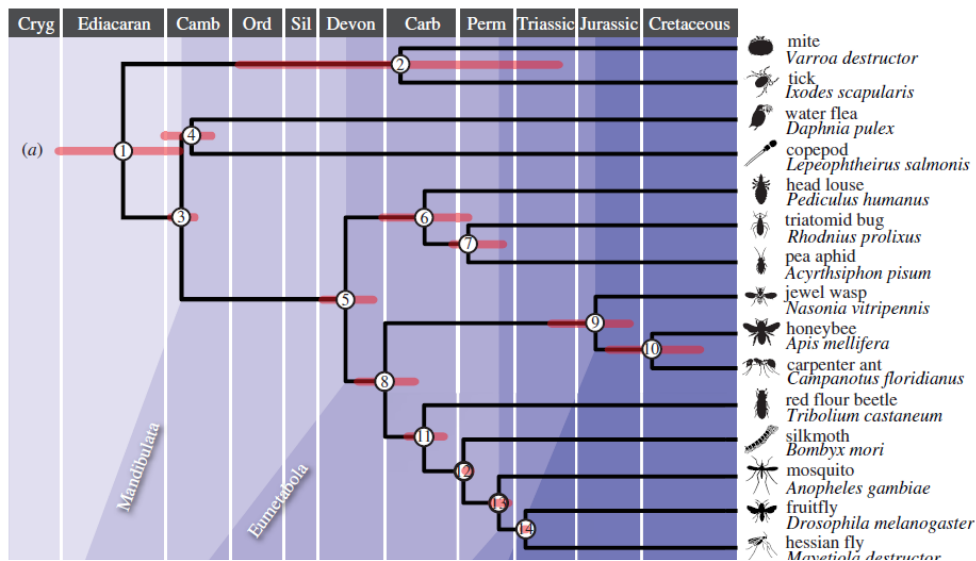
# FOSSIL CALIBRATION



$\geq 161$  Ma

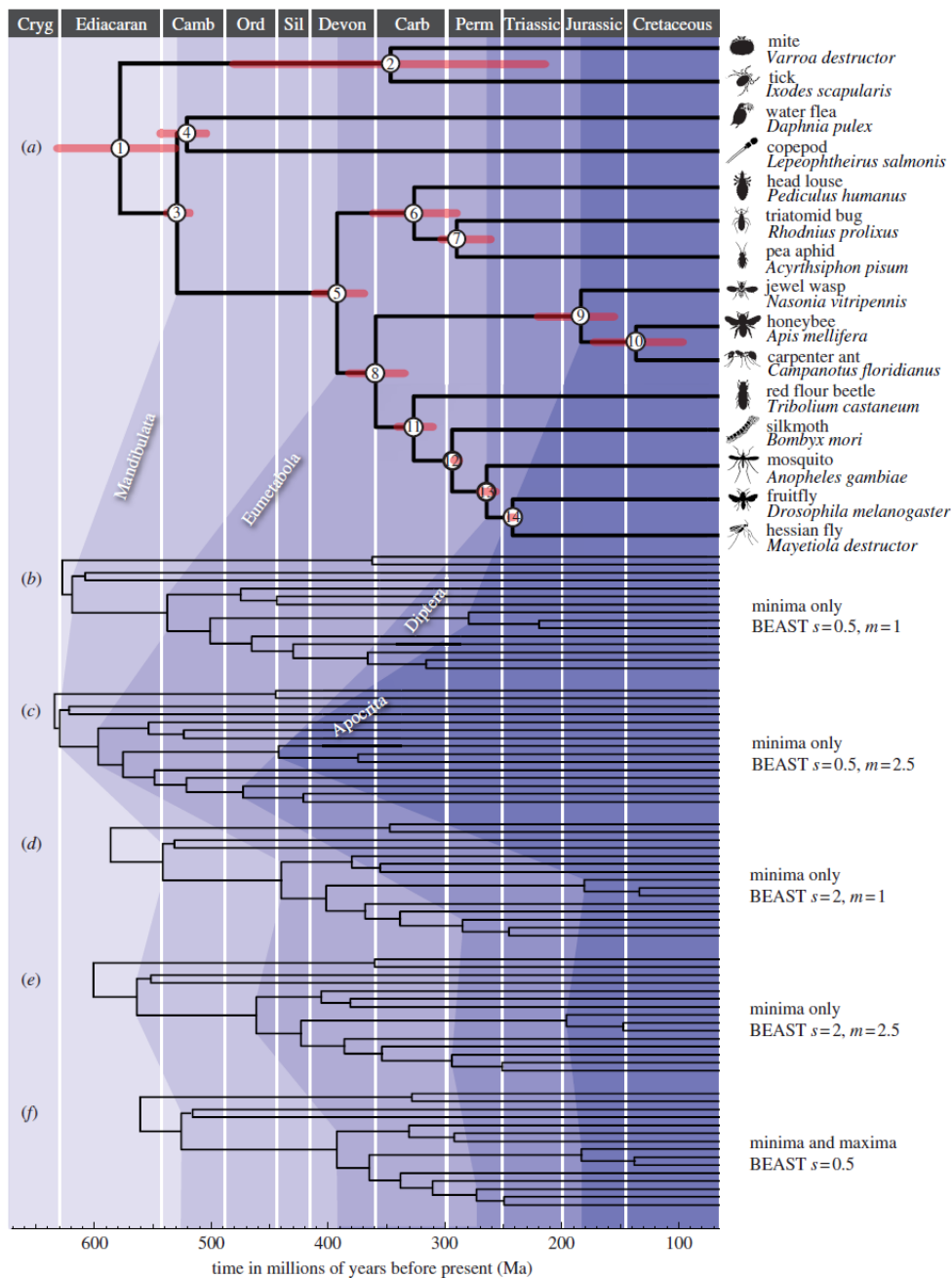


hard **minimum** = fossil age  
**mean** of distribution?  
which distribution?



## Warnock, Yang & Donoghue 2011, Biol. Lett.

- sensitivity to priors on calibration points
- uniform prior
- lognormal prior with different mean and std deviation



## Warnock, Yang & Donoghue 2011, Biol. Lett.

- sensitivity to priors on calibration points
- uniform prior
- lognormal prior with different mean and std deviation

# BIOGEOGRAPHIC CALIBRATION

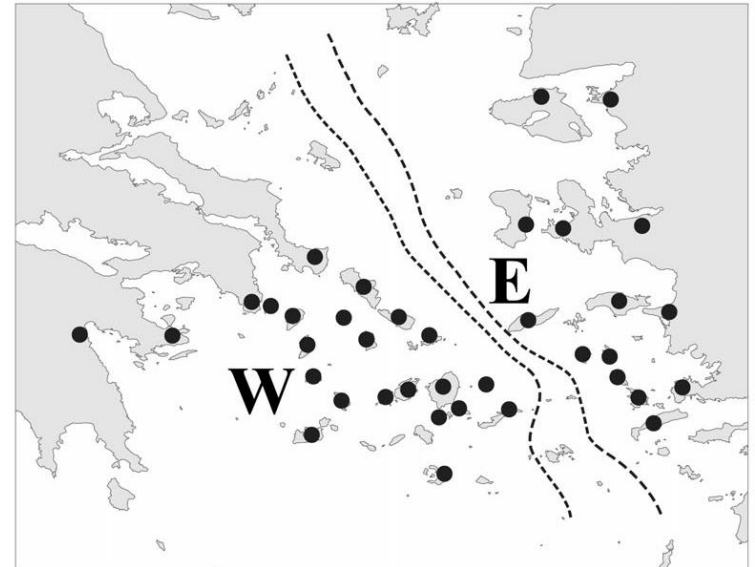
## **biogeography**, e.g. continental breakups

- associate a split on phylogeny with a geographic event
- corresponding node calibrated to the estimated time of this event (or time range)
- gives minimum **and** maximum

# BIOGEOGRAPHIC CALIBRATION

## Papadopoulou et al. 2010, MBE

- Aegean sea, mid-aegean trench (9-12 my)
  - flightless beetles
  - eight beetle lineages split into West and East group
- CO1 substitution rate of 3.5% per my



# BIOGEOGRAPHIC CALIBRATION

## Songbirds (Passeriformes)

- New Zealand wrens sister to remaining songbirds
- NZ split off gradually 52-85 Mya
- vicariance?
- Ericson et al. 2014: large impact!





# SECONDARY CALIBRATION

## using results from previous dating studies

- use posterior 95% range from a previous study as a calibration point
- either with normal or uniform distribution
- danger of propagating error... – and not adding much to the previous results!

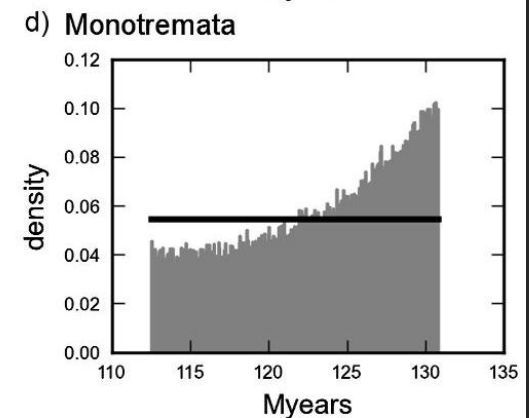
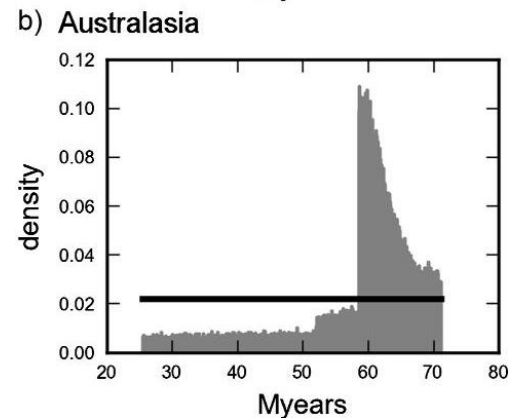
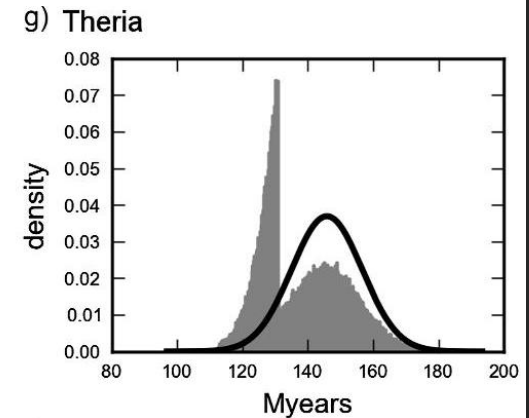
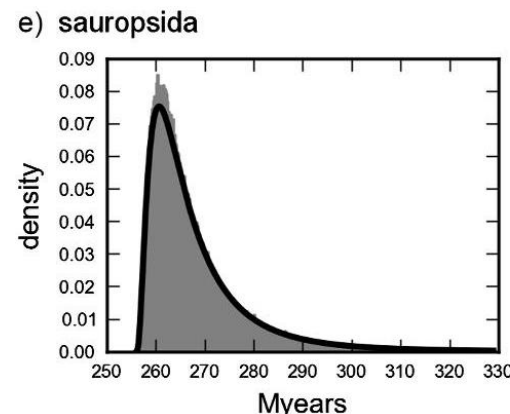
# NODE CALIBRATION

## multiple calibrations

- interactions of node age priors!
- effective priors even on uncalibrated nodes

→ run without data

Heled & Drummond 2012



# NODE CALIBRATION

**analysis is as good as your prior!**

- assess realized node calibration priors
- test sensitivity to age priors
- investigate alternative interpretations (of fossils, biogeography, previous studies...)

# NODE CALIBRATION

## Dinosaurs didn't have fleas

“... we estimated divergence dates with a validated set of 37 fossils.”

supplementary:

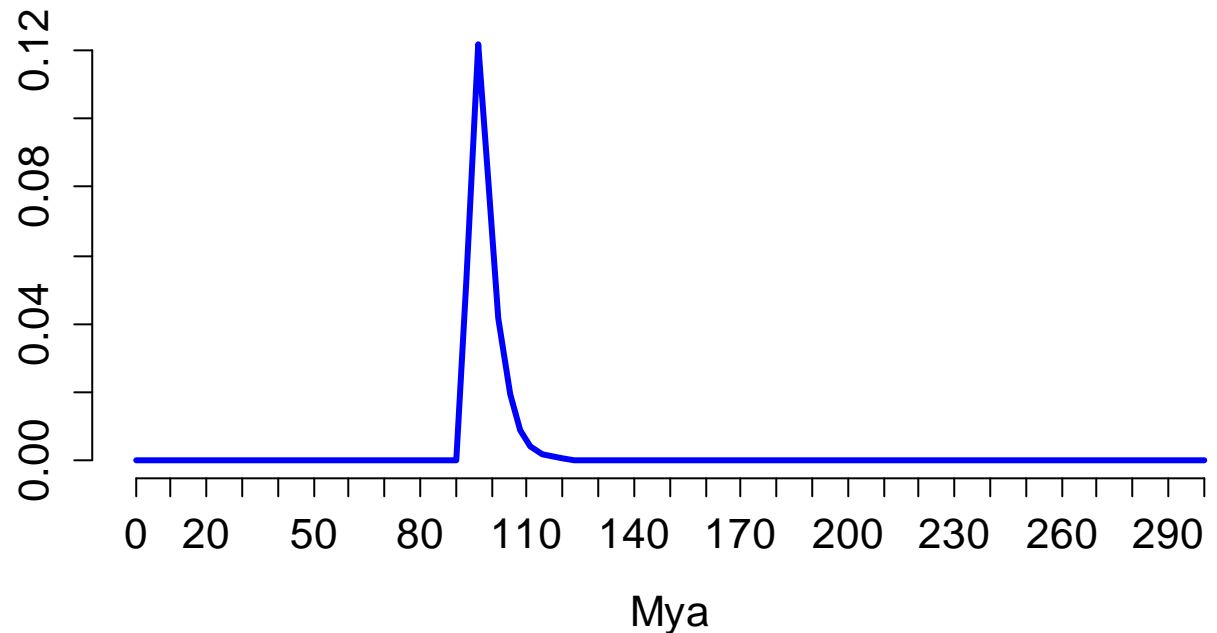
- 37 fossils, careful placement, for minima
- maxima: sufficient land masses
- 'log-normal' ( $\mu = 2$ ;  $d = 0.5$ ) distribution



# NODE CALIBRATION

**Dinosaurs didn't have fleas**

example: calibration on ant-bee divergence

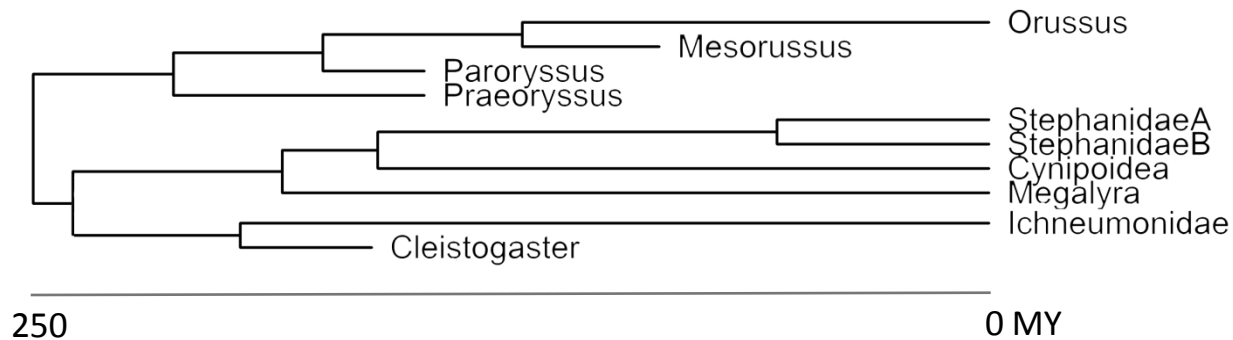


# TIP CALIBRATION

## tip dating

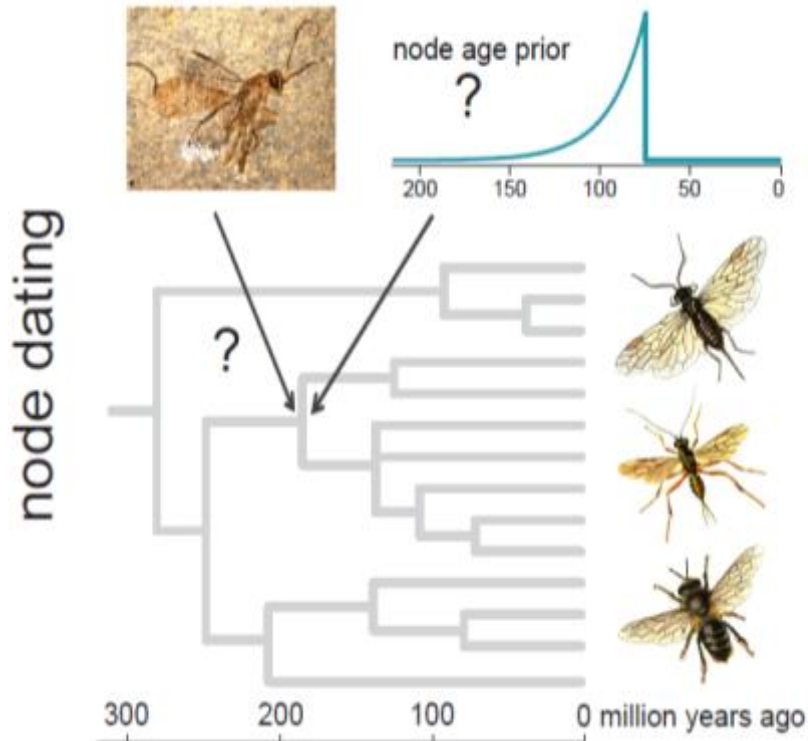
- calibrating the tree by tips that are sampled through time
- various data sources
  - viruses: sampled in different years
  - using ancient DNA: samples from different times
  - fossils: treat morphology as original data (scored for fossils and extant taxa alike)

# TIP CALIBRATION

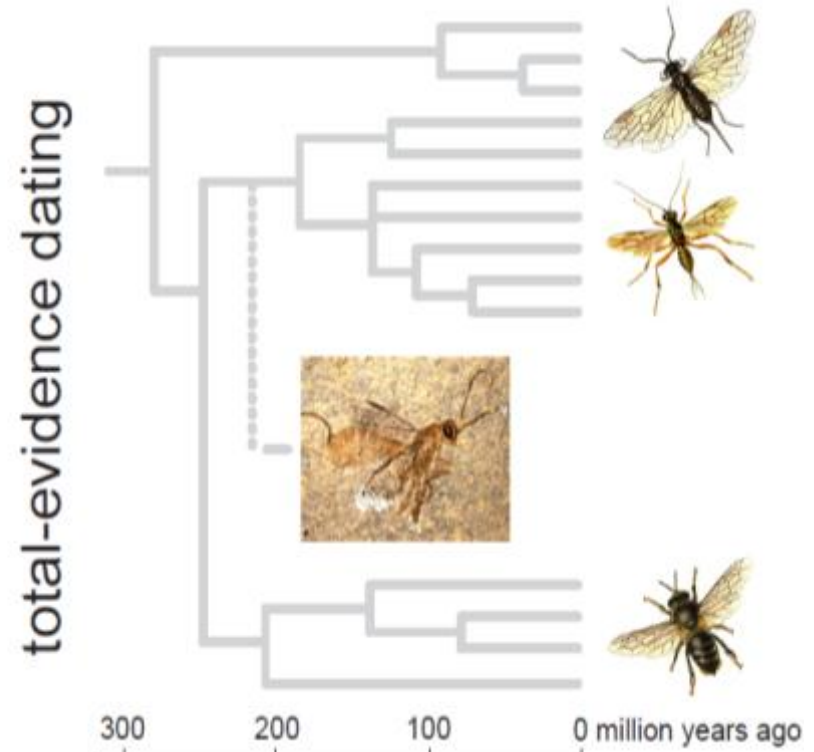


- trees with terminals at different times
- no external calibration needed!
- most elegant way, fewest assumptions

# TIP CALIBRATION



fossils used to derive  
priors on node ages



fossils used as primary  
observations (morphology)

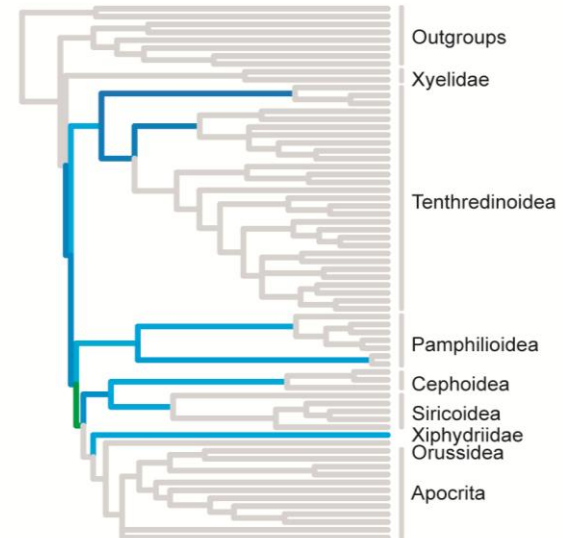
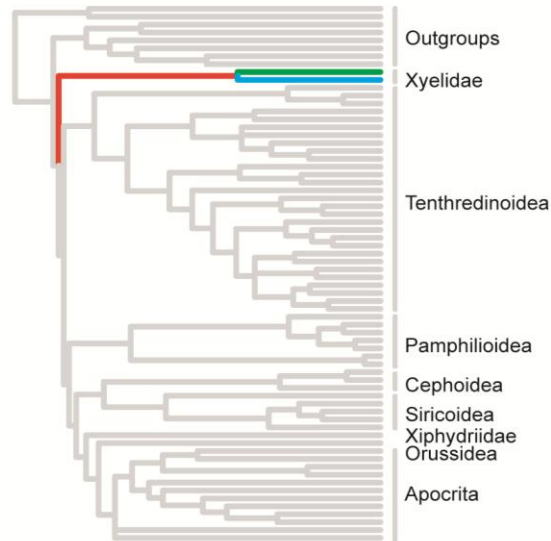
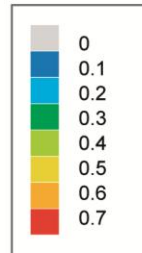


# TIP CALIBRATION

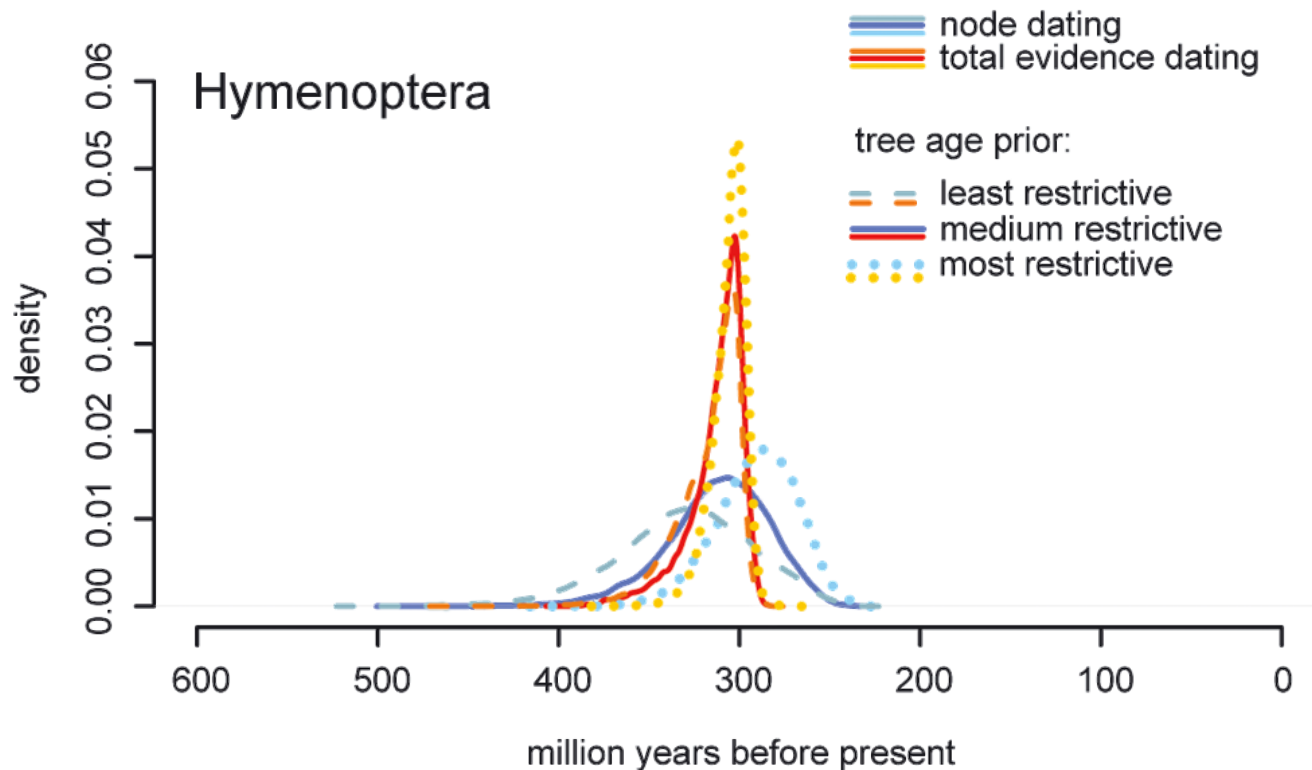
- **explicitly** includes fossil evidence (as data, not as prior assumptions)
- capable of including **all** fossil evidence
- no need for *a priori* placement of fossils, instead **integrates over the uncertainty** in fossil placement
- assumes morphological clock

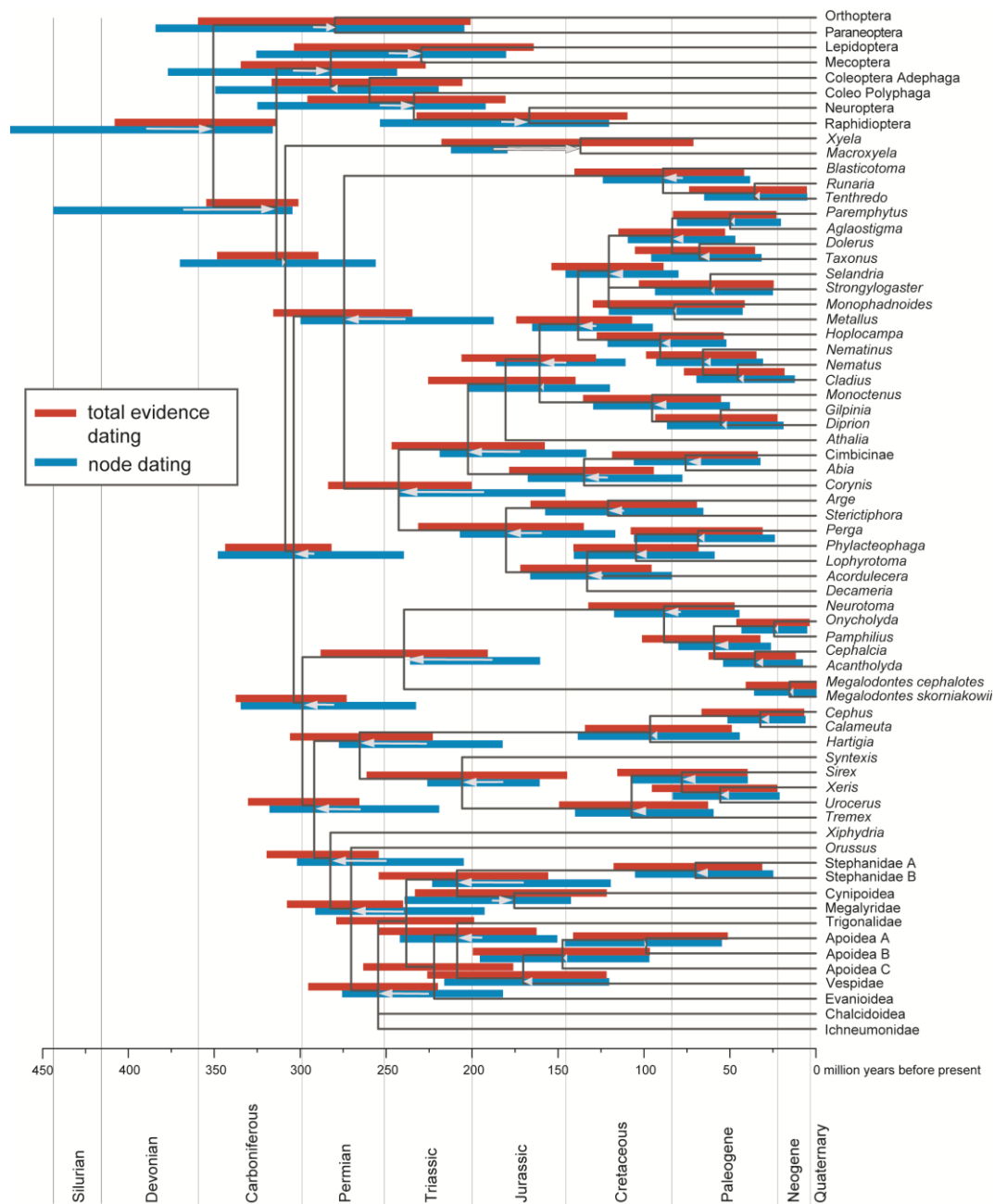
# TIP CALIBRATION: HYMENOPTERA

uncertain fossil placement



# TIP CALIBRATION: HYMENOPTERA





- usually higher precision with TE
- older ages in Hymenoptera

# CALIBRATING THE CLOCK

- most vital step of any dating analysis
- most prone to error
- most neglected



Opinion

*TRENDS in Genetics* Vol.20 No.2 February 2004

Full text provided by [www.sciencedirect.com](http://www.sciencedirect.com)



## Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision

Dan Graur<sup>1</sup> and William Martin<sup>2</sup>

<sup>1</sup>Department of Biology and Biochemistry, University of Houston, Houston, TX 77204-5001, USA

<sup>2</sup>Institut für Botanik III, Heinrich-Heine Universität Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany