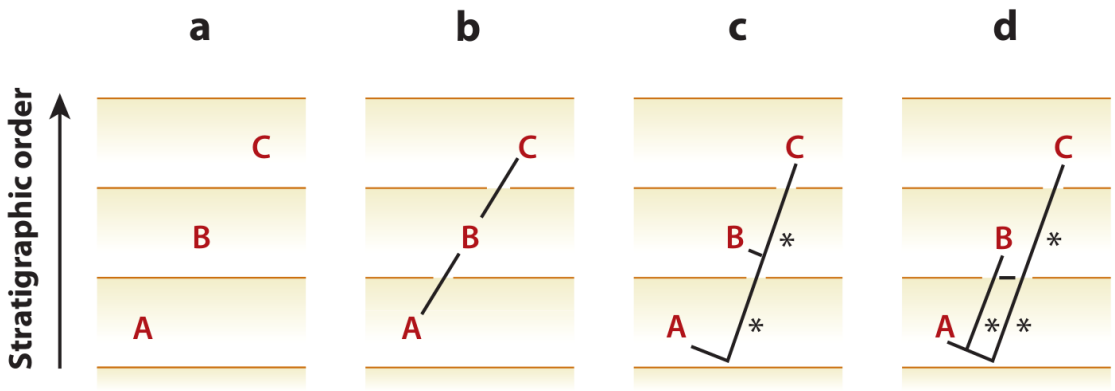
Employing models of stratigraphic preservation in phylogenetic inference

Caroline Parins-Fukuchi, Daniel C. Fisher
University of Michigan

Stratocladistic approach

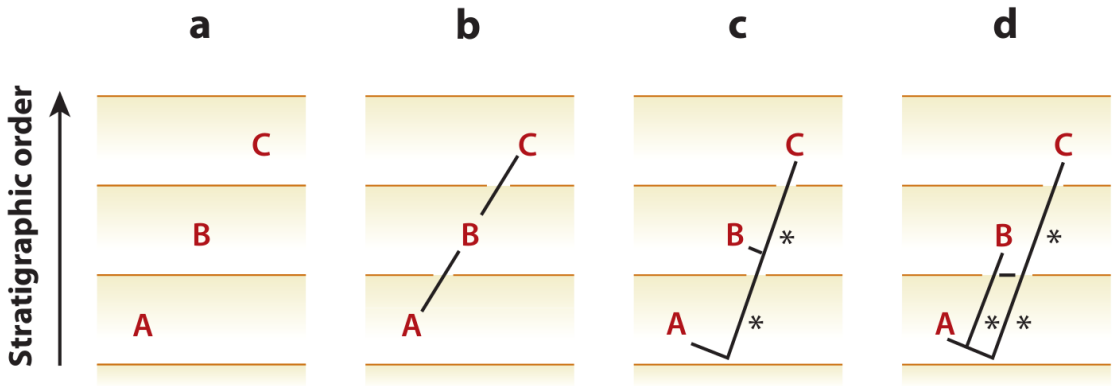
Temporal data can inform phylogenetic relatedness



Fisher 2008

Stratocladistic approach

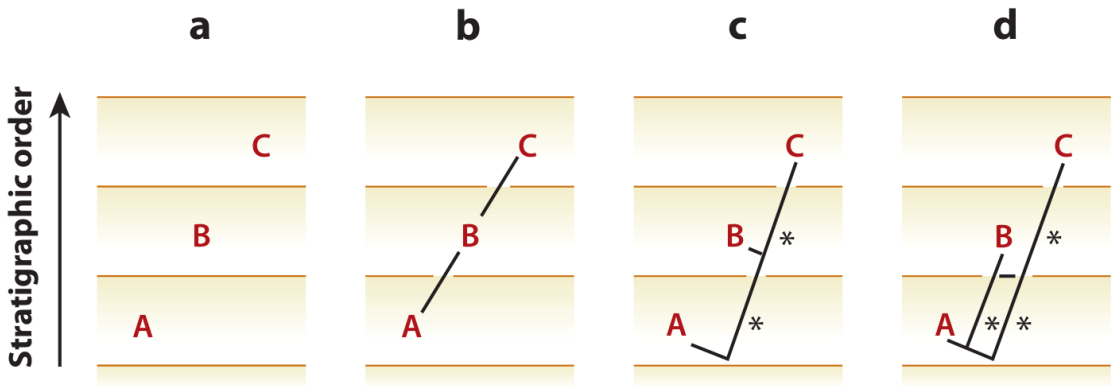
Choose topology which implies fewest unsampled stratigraphic crossings ("stratigraphic debt")



Fisher 2008

Stratocladistic approach

- Relies on parsimony principal
- Combine with character data to minimize combined parsimony debt

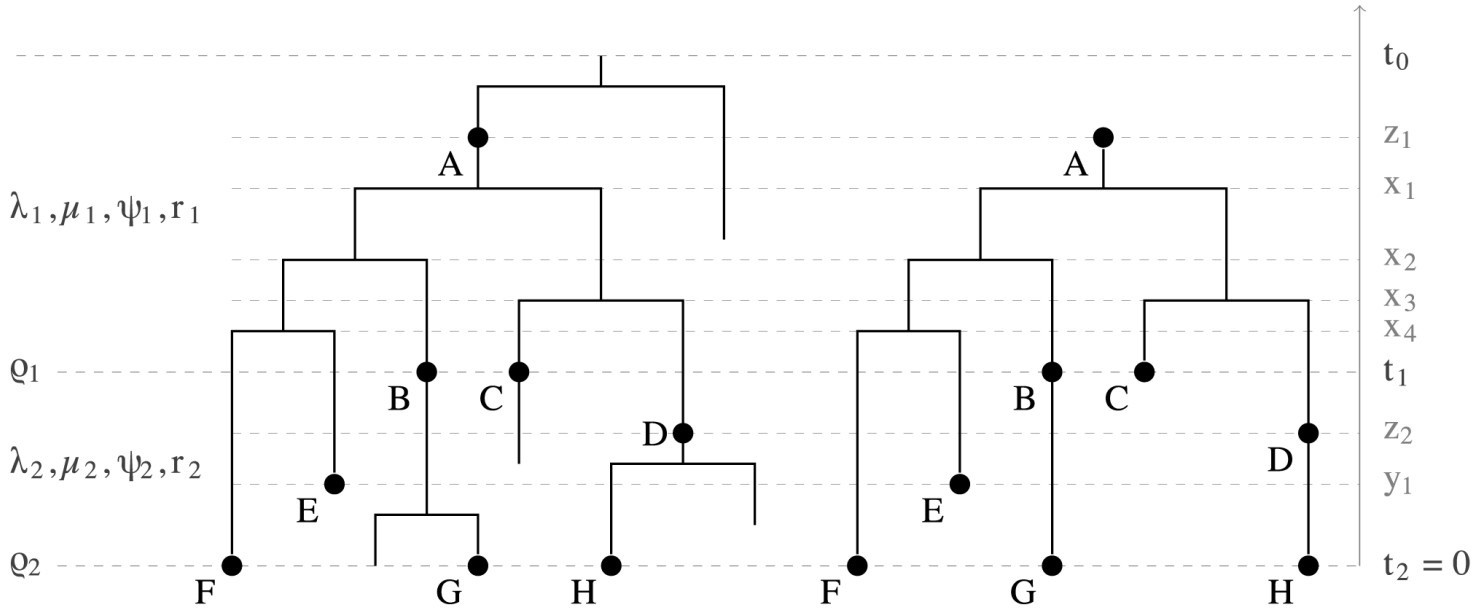


Fisher 2008

Stratocladistic likelihood

Could be applied in a probabilistic modelling framework:

- Evaluate tree likelihood under models of fossil preservation
- Compare different models
- Explicit statement of model assumptions



Bayesian tip dating, especially combined with Fossilized Birth-Death (FBD) prior could be used to address similar questions

- Will not discuss here
- Instead, apply preservation and character models in maximium likelihood (ML) framework

Stratocladistic likelihood

- Calculate likelihood as product of independent Poisson processes along b branches:

$$\hat{L} = \prod_{i=1}^{b} \frac{(o_s^i - o_f^i)^{n_o^i - 2} \lambda^{n_o^i} e^{-\lambda(t_s^i - t_f^i)}}{(n_o^i - 2)!}$$

(Huelsenbeck and Rannala 1997)

Stratocladistic likelihood

Characteristics of model:

- Assumes preservation rate is constant across lineages
- Branching time estimates will match first and last fossil deposit times as closely as is permitted by topology
  - Best thought of as minimum divergence times
- Likelihood of branch lengths asymptotes as they approach observed range

$$\hat{L} = \prod_{i=1}^{b} \frac{(o_s^i - o_f^i)^{n_o^i - 2} \lambda^{n_o^i} e^{-\lambda(t_s^i - t_f^i)}}{(n_o^i - 2)!}$$

Stratocladistic likelihood

- Combine with discrete and continuous character evolutionary models (ex. Lewis Mk, Brownian motion) by summing log-likelihoods

$$\hat{L} = \prod_{i=1}^{b} \frac{(o_s^i - o_f^i)^{n_o^i - 2} \lambda^{n_o^i} e^{-\lambda(t_s^i - t_f^i)}}{(n_o^i - 2)!}$$

Stratocladistic likelihood

- Combine with discrete and continuous character evolutionary models (ex. Lewis Mk, Brownian motion) by summing log-likelihoods
- **Test hypotheses of direct ancestorship/anagenetic evolution by collapsing branches and comparing likelihood scores**

$$\hat{L} = \prod_{i=1}^{b} \frac{(o_s^i - o_f^i)^{n_o^i - 2} \lambda^{n_o^i} e^{-\lambda(t_s^i - t_f^i)}}{(n_o^i - 2)!}$$

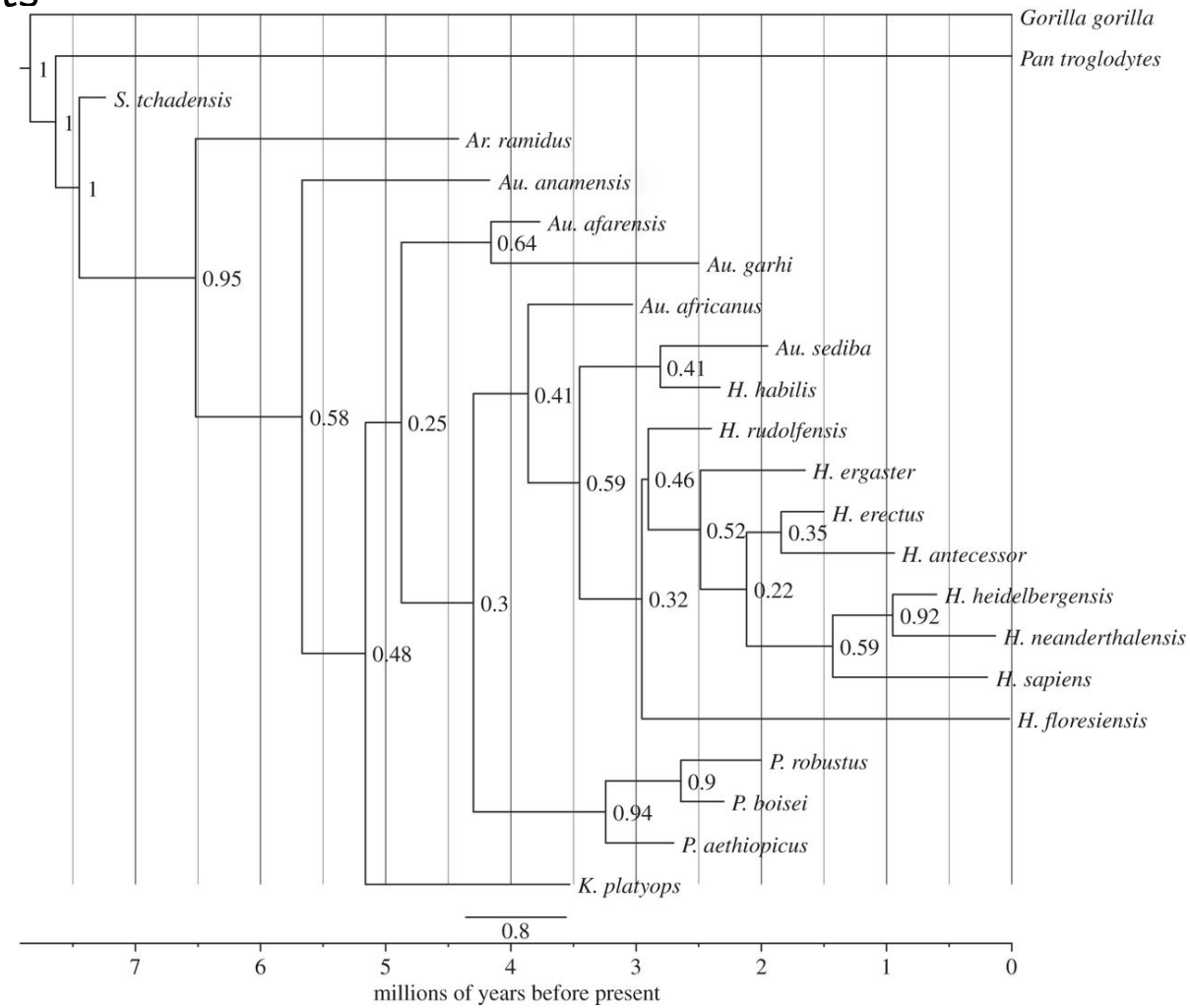
Mandos

- Implemented in 'mandos' package (github.com/carolinetomo/mandos)– working, needs optimisation
- Calculate stratigraphic and morphological likelihoods on candidate trees
  - Likelihoods calculated using Lewis Mk, Brownian motion, and Poisson preservation models
- Test direct ancestor-descendant hypotheses in likelihood framework
  - Compare likelihoods (AIC) of collapsed and bifurcating arrangements

Direct Ancestors in the Hominin Fossil Record

Application to fossil hominins

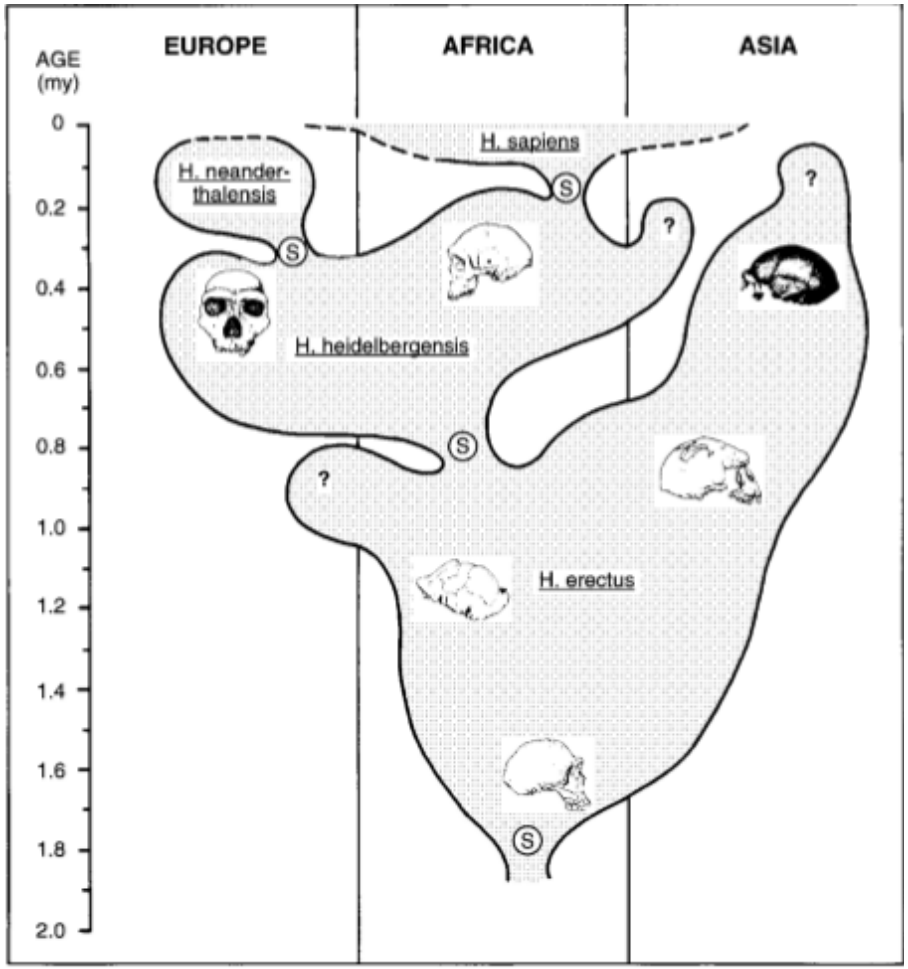
Trees estimated from morphology alone disagree with paleoanthropological assessments



Dembo et al. 2015

Application to fossil hominins

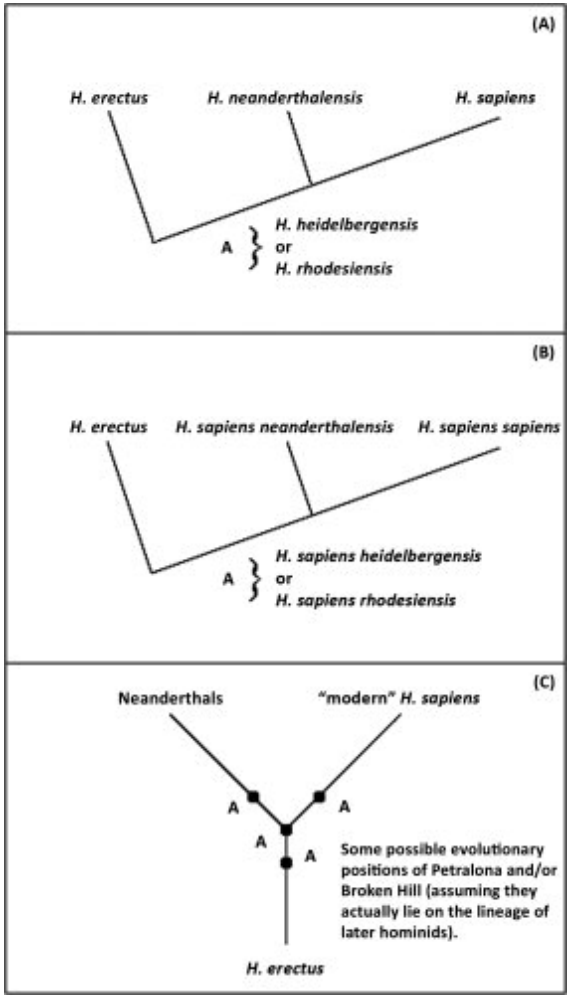
Paleoanthropological hypotheses often consider direct-ancestorship



Rightmire 1997

Application to fossil hominins

Paleoanthropological hypotheses often consider direct-ancestorship



Stringer 2012

Fossil hominins

- Published morphological matrix (Dembo et al. 2015)
  - 23 taxa, 391 characters
- Temporal and stratigraphic occurrences surveyed across literature

Analysis

Generate 1700 candidate topologies from morphological ML/parsimony starting tree

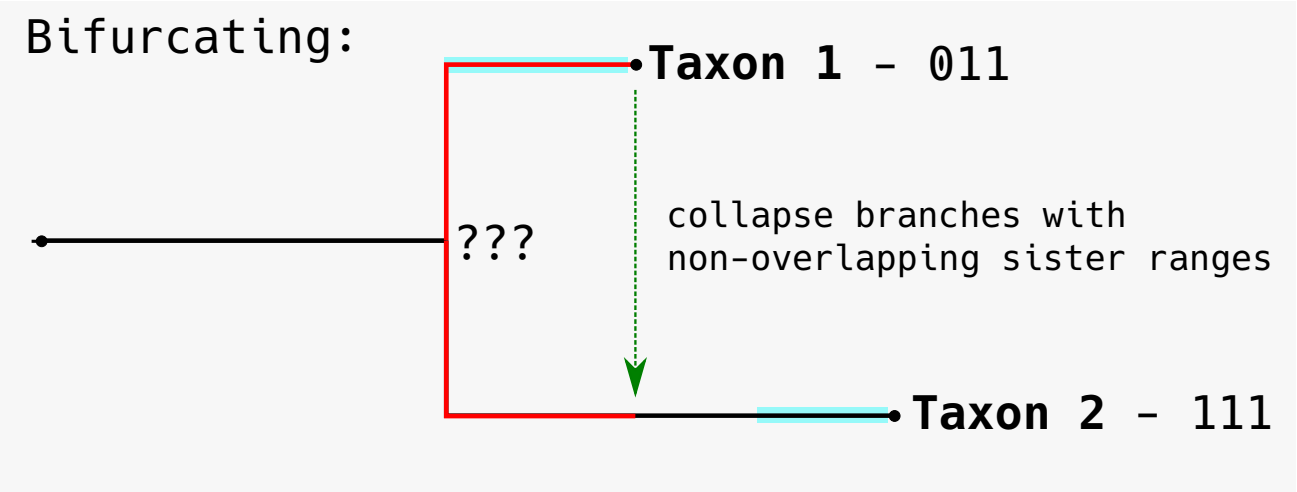
Calculate likelihood of stratigraphic and morphological data on each topology

Collapse individual branches and test for anagenesis on best tree

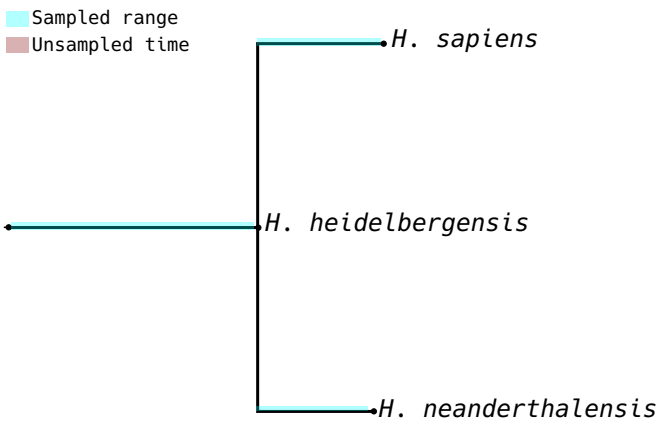
Branch support measured as approximate posterior probabilities

Testing for direct ancestorship

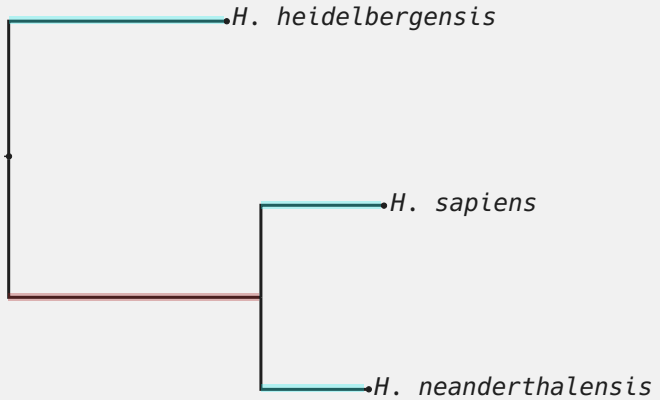
Branches with non-overlapping sister ranges are collapsed as set of potential direct ancestors



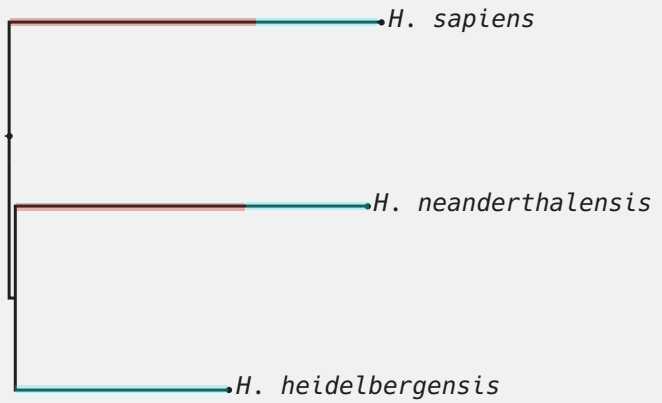
Fossil hominins



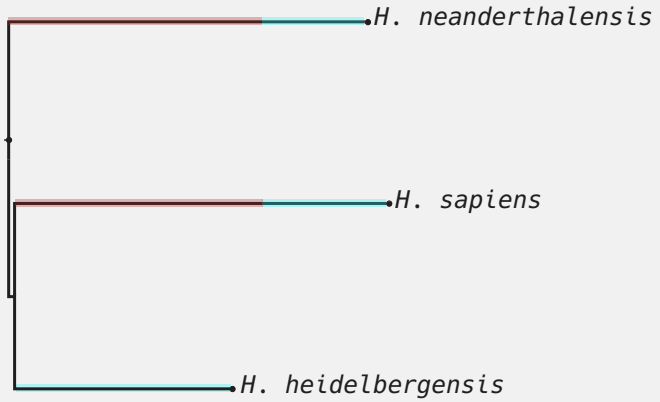
Mk log-likelihood: **-1703.00**
Stratigraphic log-likelihood: **-52.16**
Combined log-likelihood: **-1755.79**



Mk log-likelihood: -1782.91
Stratigraphic log-likelihood: -67.87
Combined log-likelihood: -1850.78

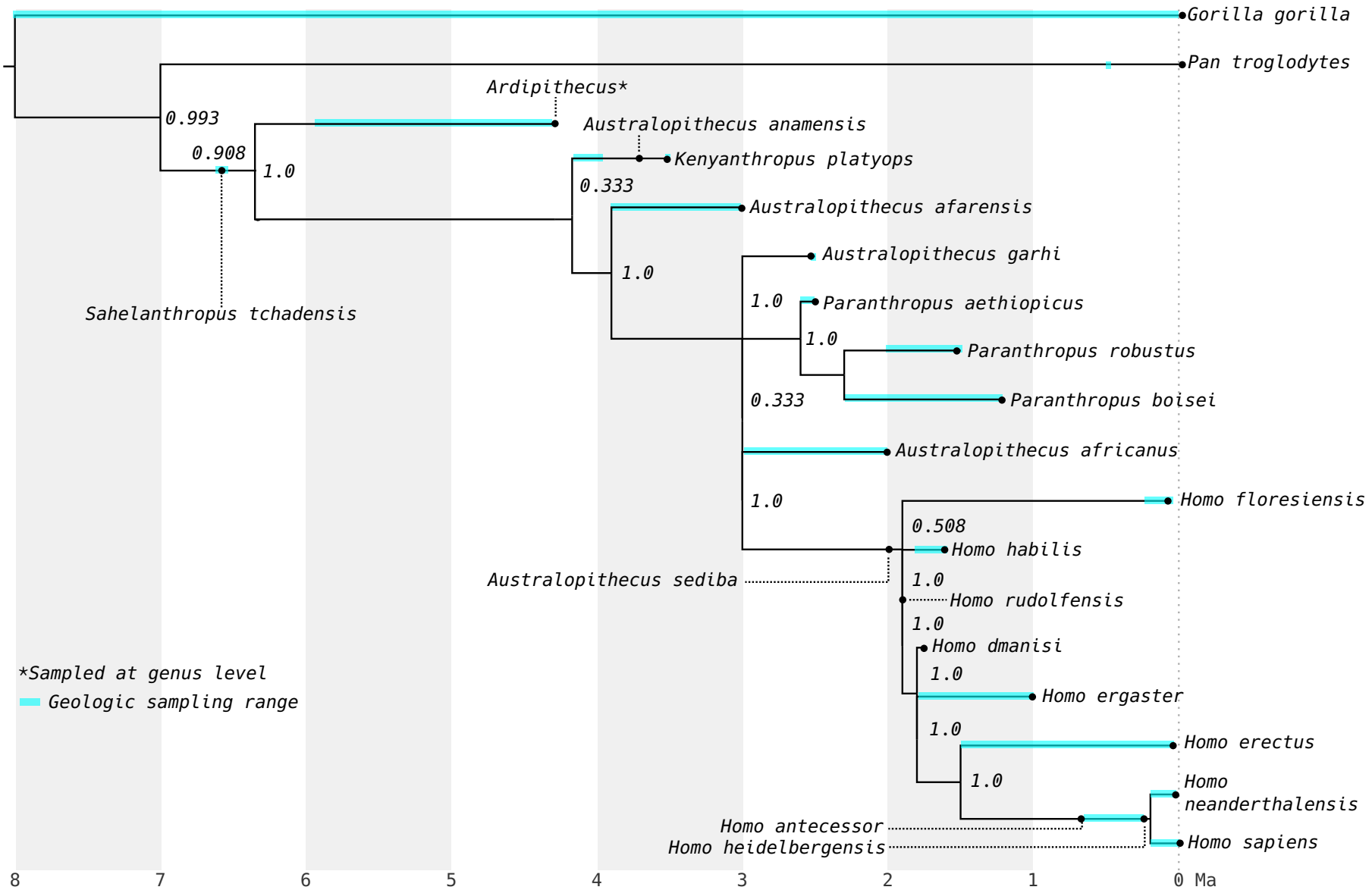


Mk log-likelihood: -1744.68
Stratigraphic log-likelihood: -77.13
Combined log-likelihood: -1821.81



Mk log-likelihood: -1727.39
Stratigraphic log-likelihood: -77.13
Combined log-likelihood: -1804.53

Fossil hominins



Acknowledgements

- Christopher Dick lab at UMich
- Laura MacLatchey (UMich Anthropology)
- Edward Ionides (UMich Statistics)
- Joseph Brown (UMich)