

Primary endosymbiosis events date to the later Proterozoic with cross-calibrated phylogenetic dating of duplicated ATPase proteins

Nicholas J. Matzke^{*1,2}, Patrick Shih¹



*matzke@nimbios.org



My QR code to get contact info, CV, website, R code! Read with e.g. Bakodo

¹ Then: CTEG: Center for Theoretical and Evolutionary Genomics
Department of Integrative Biology, University of California, Berkeley

² Now: Postdoctoral Fellow, National Institute of Mathematical and Biological Science (NIMBioS)
and Department of Ecology & Evolutionary Biology, University of Tennessee, Knoxville

Introduction

Chloroplasts and mitochondria descended from bacterial ancestors, but the dating of these primary endosymbiosis events remains very uncertain, despite their importance for our understanding of the evolution of both bacteria and eukaryotes.

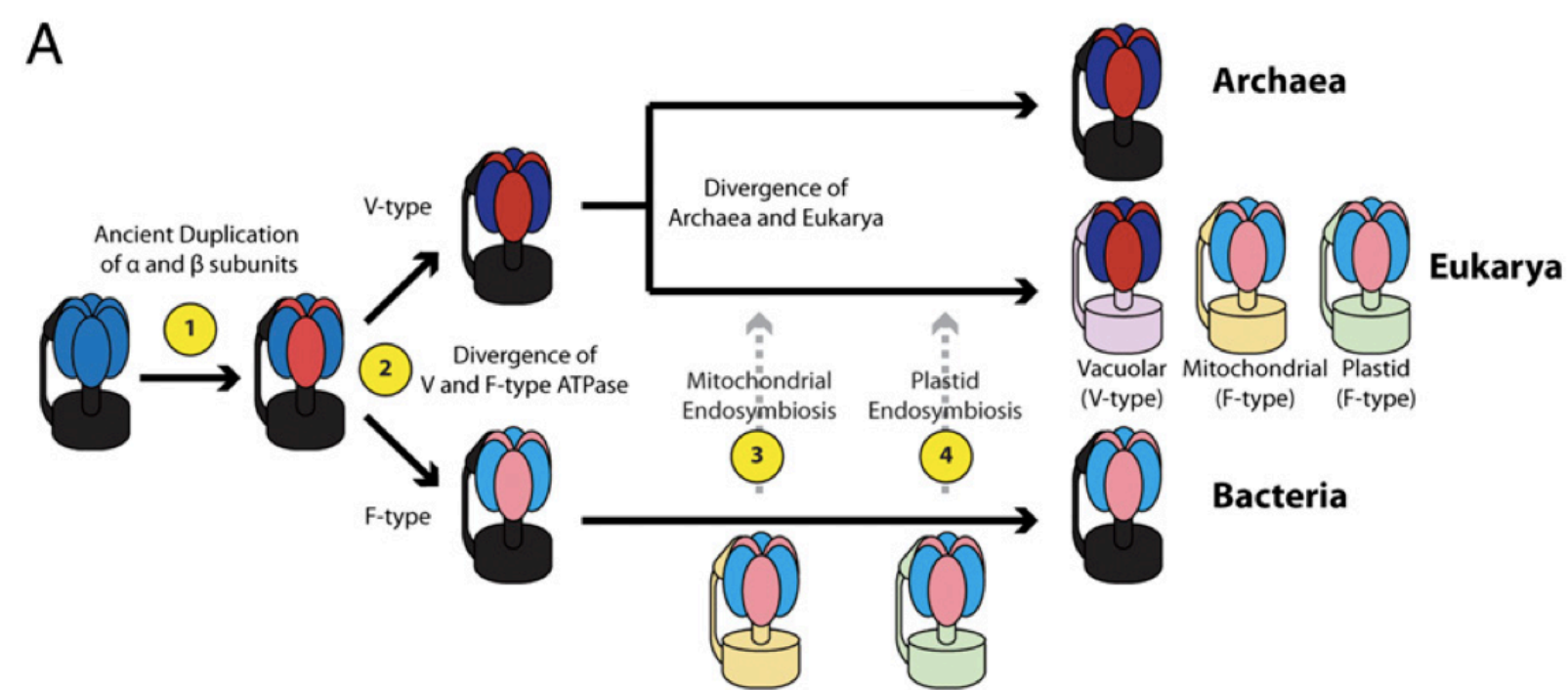
All phylogenetic dating in the Proterozoic and before is difficult: Significant debates surround potential fossil calibration points based on the interpretation of the Precambrian microbial fossil record, and strict molecular clock methods cannot be expected to yield accurate dates over such vast timescales because of strong heterogeneity in rates. Even with more sophisticated relaxed-clock analyses, nodes that are distant from fossil calibrations will have a very high uncertainty in dating.

However, endosymbiosis events and gene duplications provide some additional information that has never been exploited in dating; namely, that certain nodes on a gene tree must represent the same events, and thus must have the same or very similar dates, even if the exact date is uncertain.

Method: Exploit common dates of gene duplicates

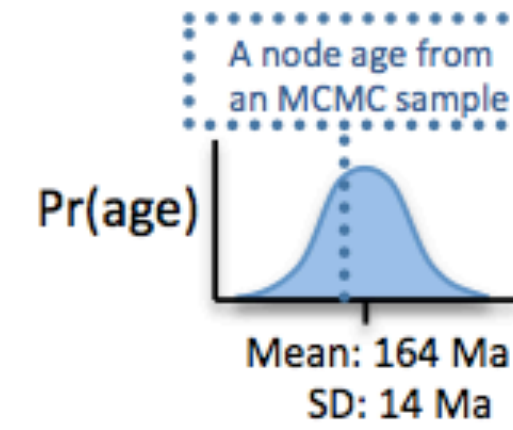
We devised techniques to exploit this information: cross-calibration, in which node date calibrations are reused across a phylogeny, and cross-bracing, in which node date calibrations are formally linked in a hierarchical Bayesian model.

We apply these methods to proteins with ancient duplications that have remained associated and originated from plastid and mitochondrial endo-symbionts: the α and β subunits of ATP synthase and its relatives, and the elongation factor thermo unstable (EF-Tu).

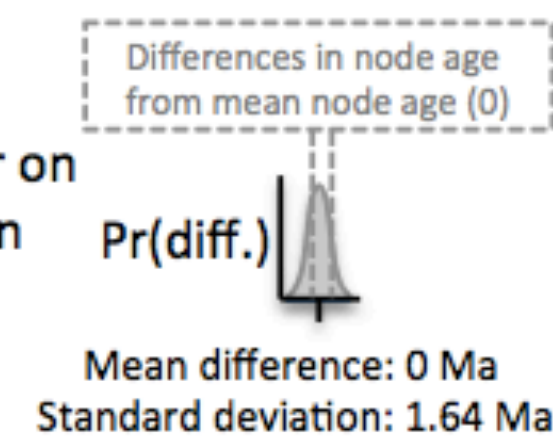


Prior distributions

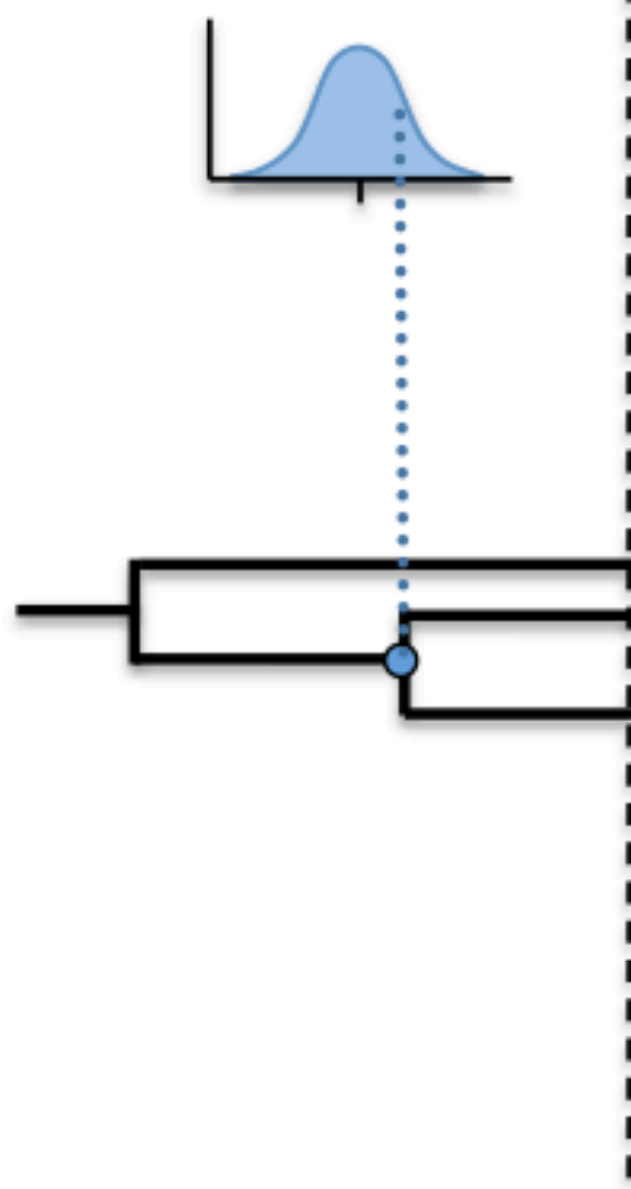
A. Traditional prior on node date, e.g. prior distribution on date of the monocot-dicot split:



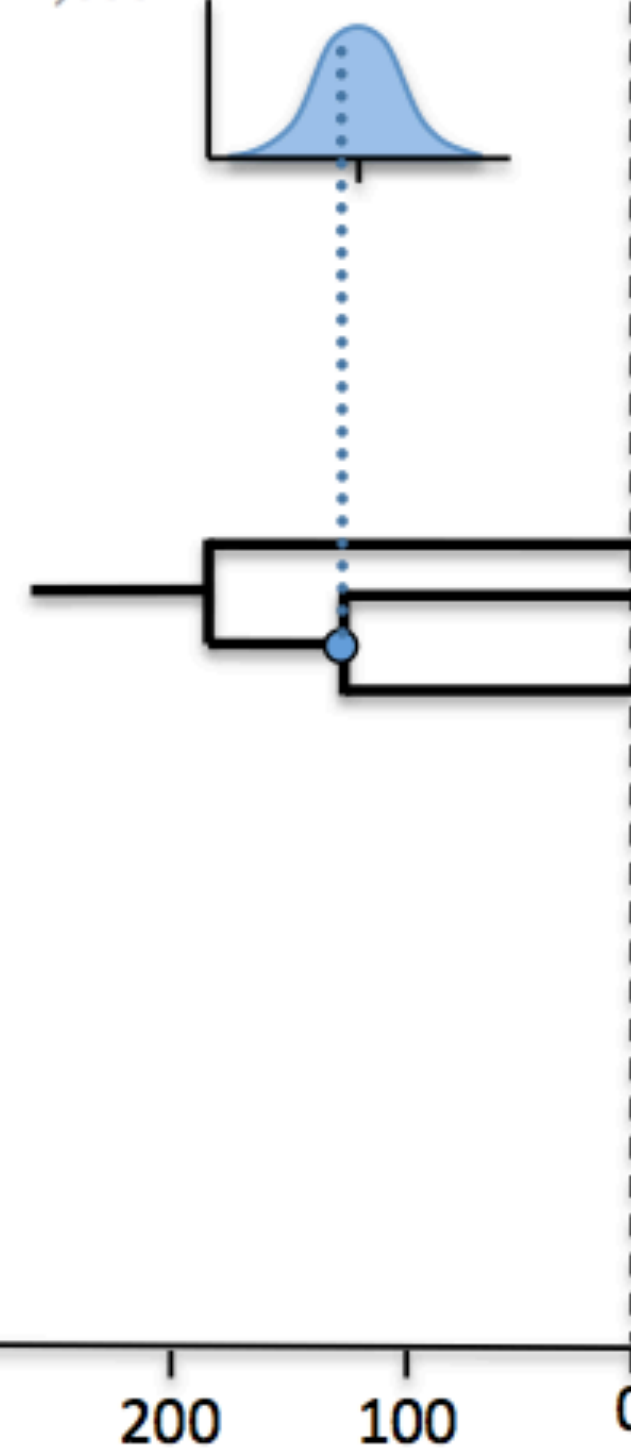
B. New prior on differences in node ages:



1. Traditional dating analysis

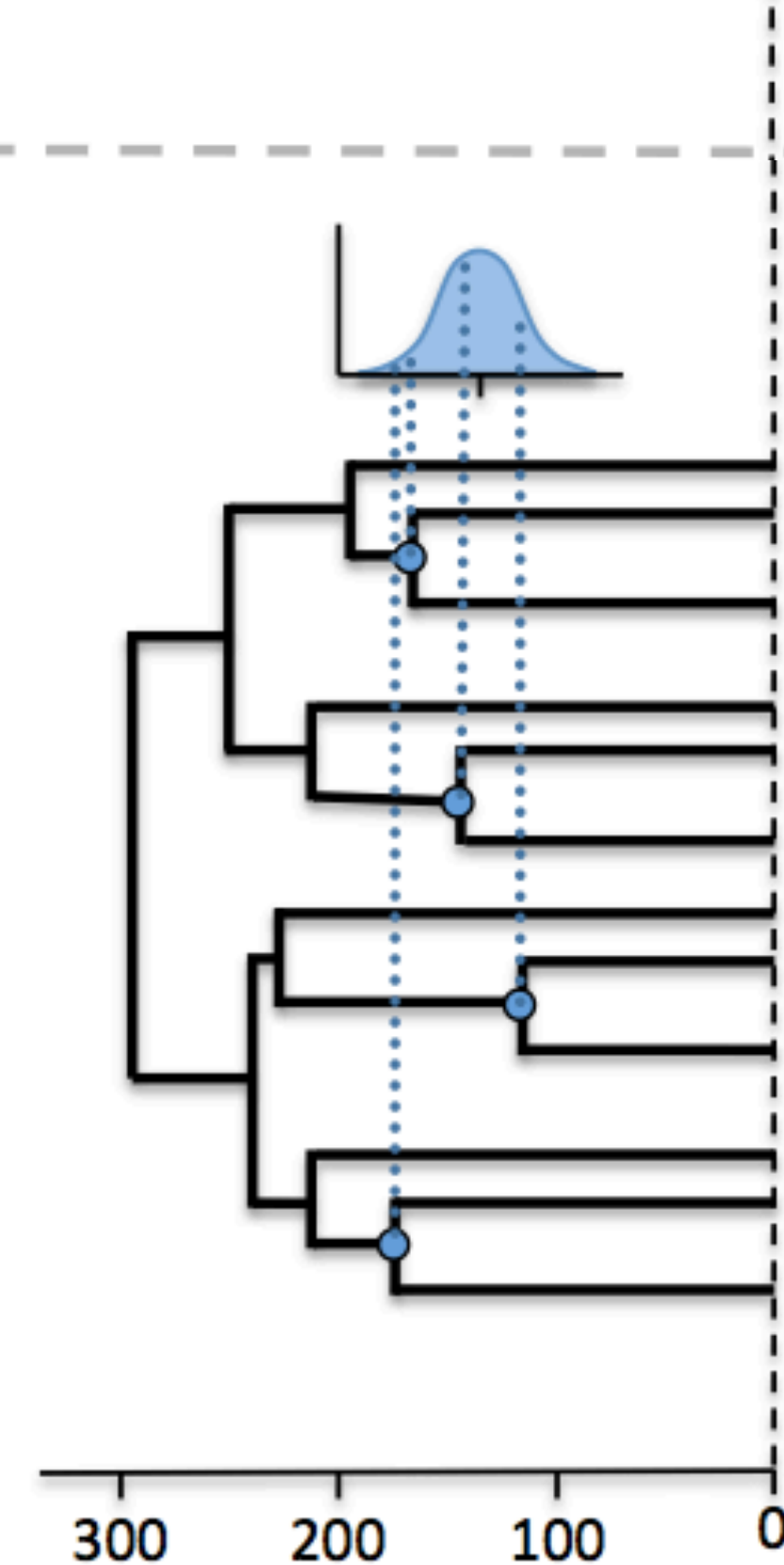
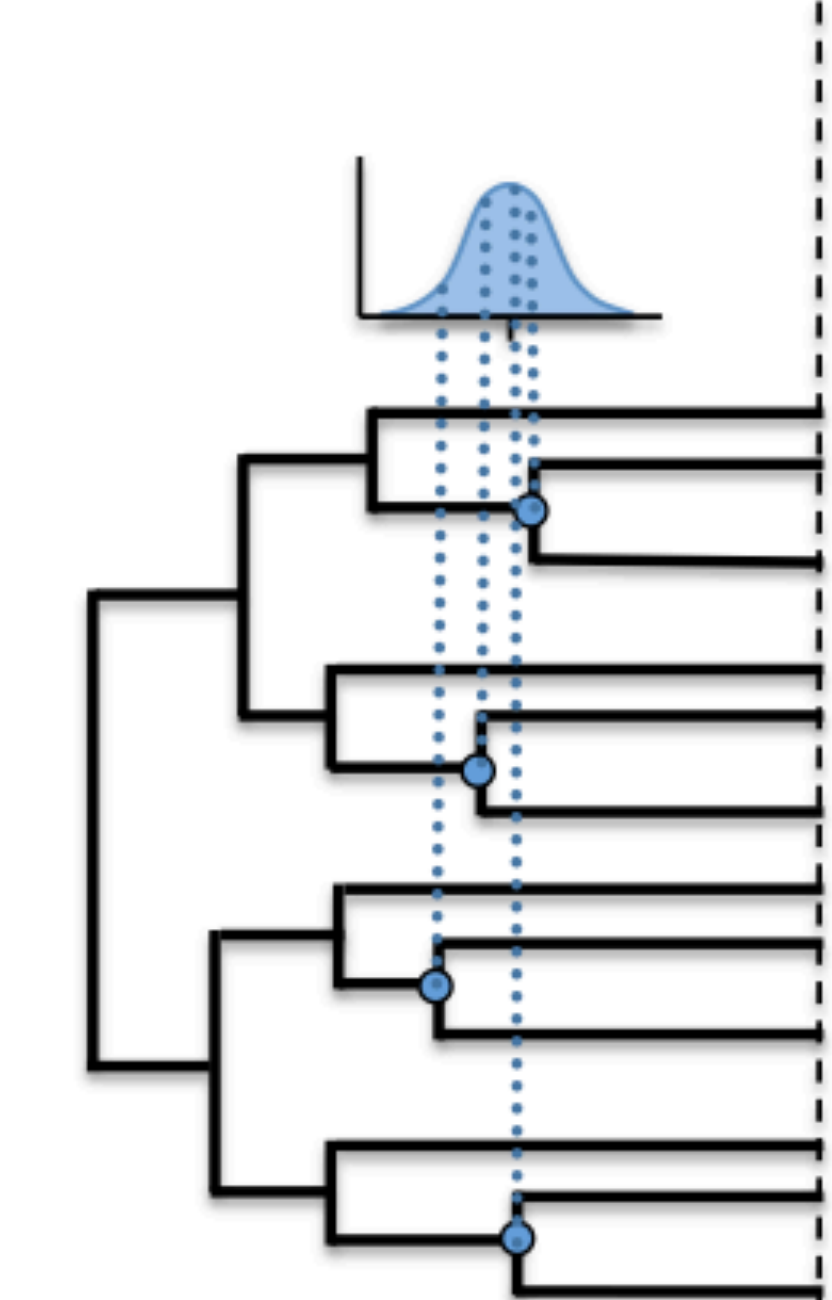


MCMC generation #: 102,000 / 112,000



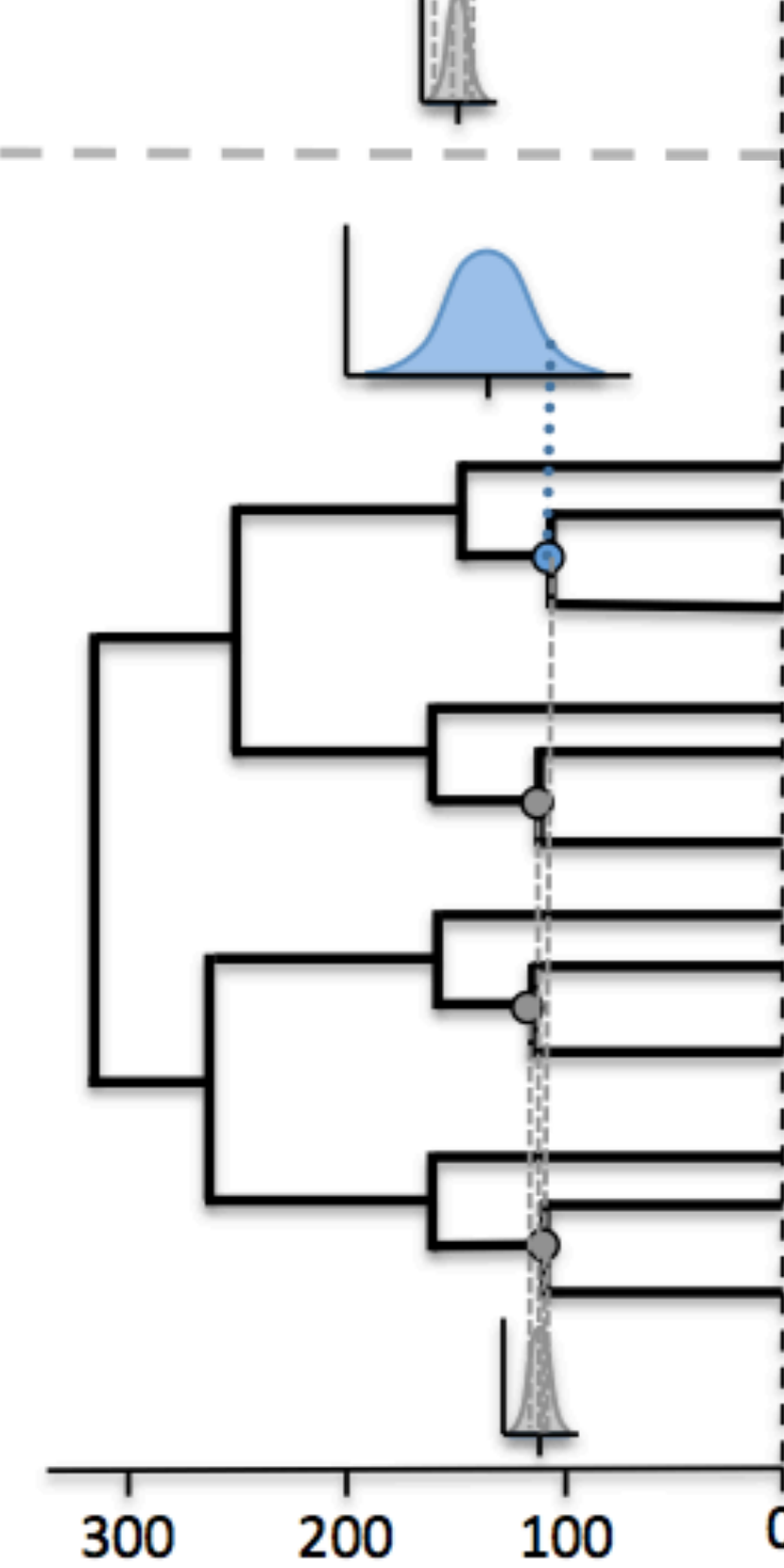
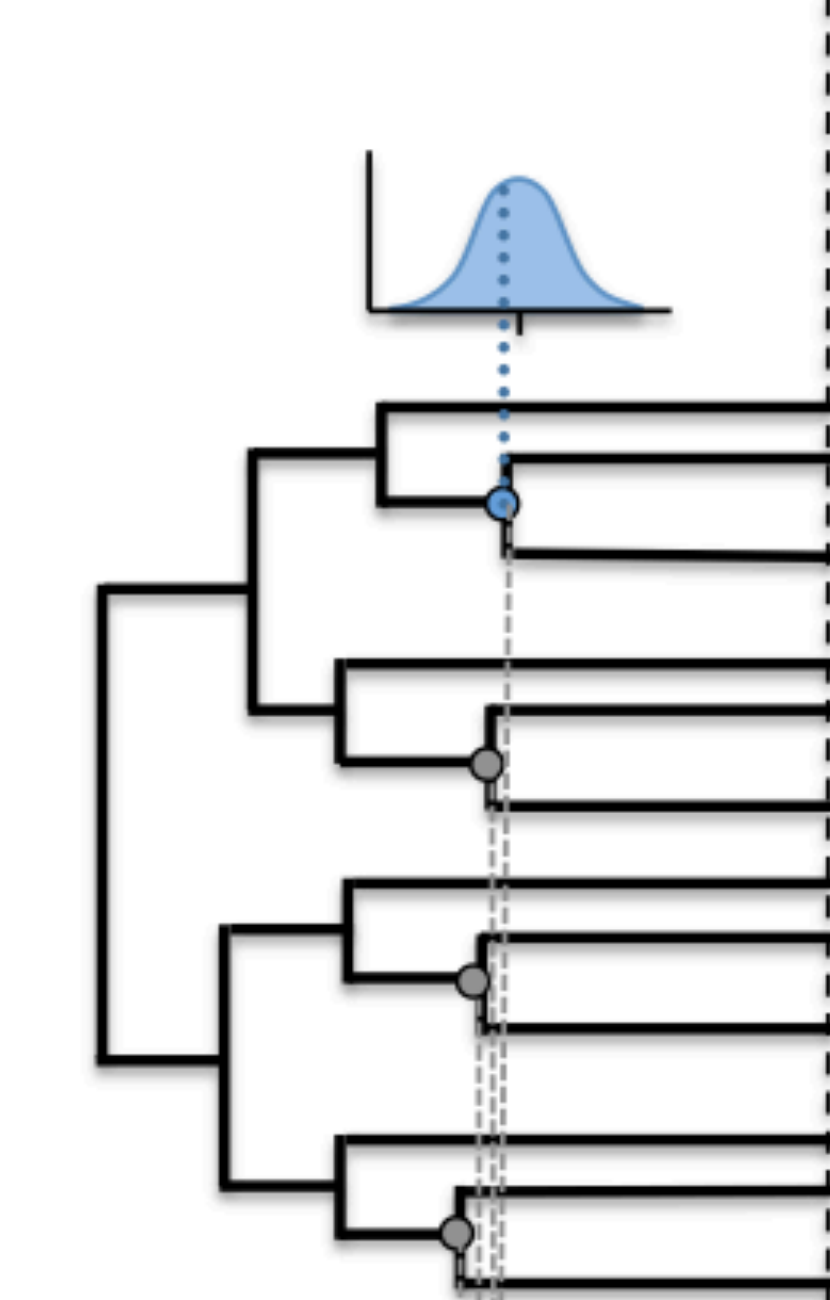
Age (m.y.a.)

2. Cross-calibration



Age (m.y.a.)

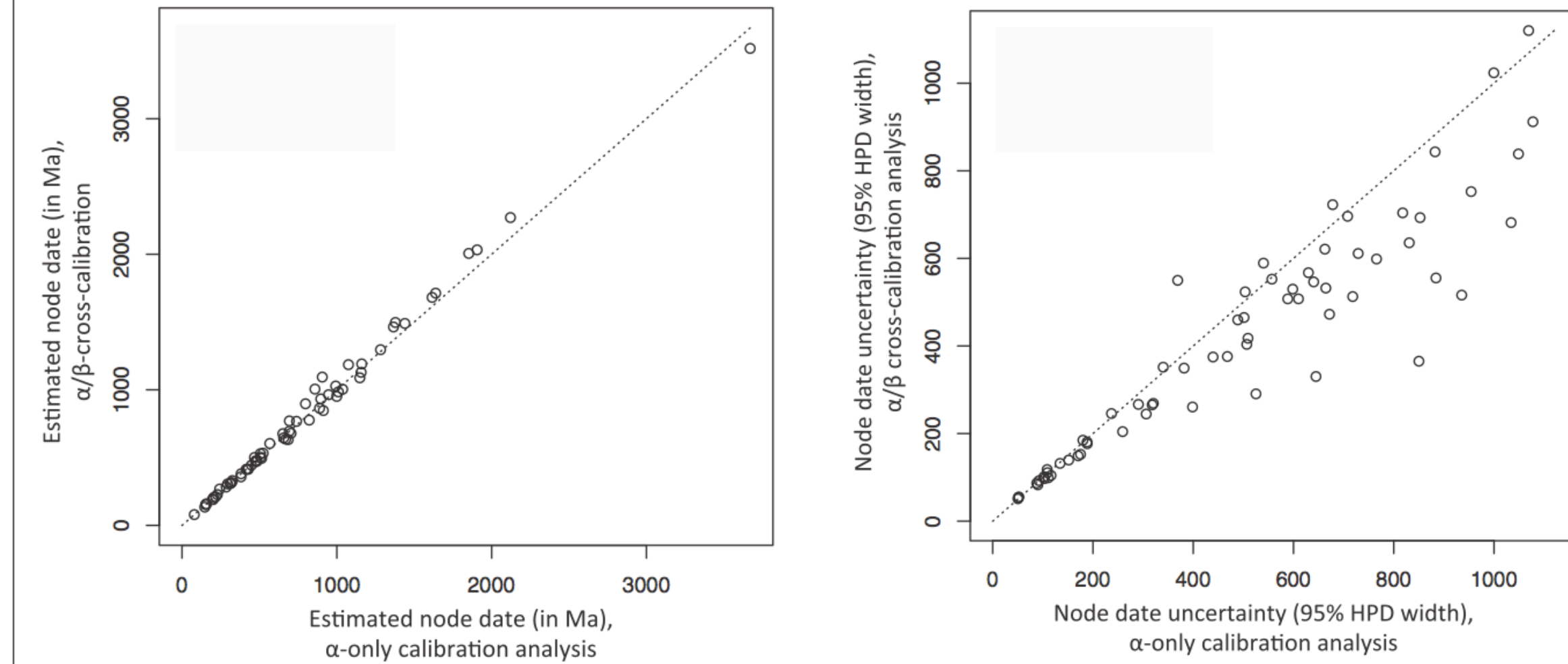
3. Cross-bracing



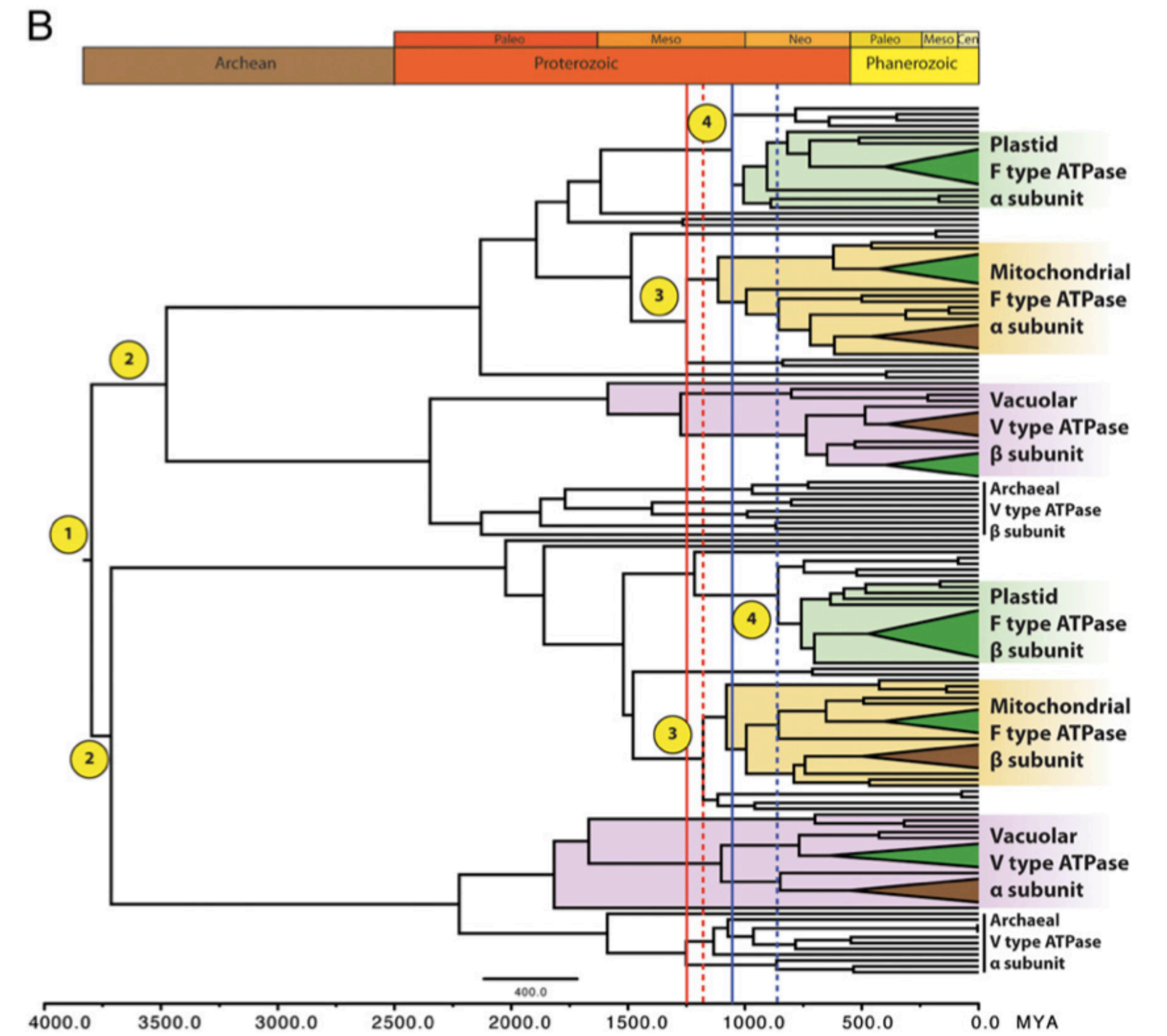
Age (m.y.a.)

Results

1. Cross-bracing doesn't shift dates. But it does reduce uncertainty.



2. The deep nesting of mitochondria and chloroplasts, and extrapolation of Phanerozoic fossil dates, suggests a late date for crown eukaryotes.



Conclusions

Hierarchical Bayesian models are very flexible and allow the incorporation of new prior information which is not exploited in traditional analyses. For this paper, see: Shih & Matzke 2013, *PNAS*, and writeup by Olga Zhaxybayeva (2013)

References & more info

Use QR code or go to: <http://phylo.wikidot.com/nicholas-j-matzke>