Lecture 2.3

Molecular Dating

Evolutionary timescales
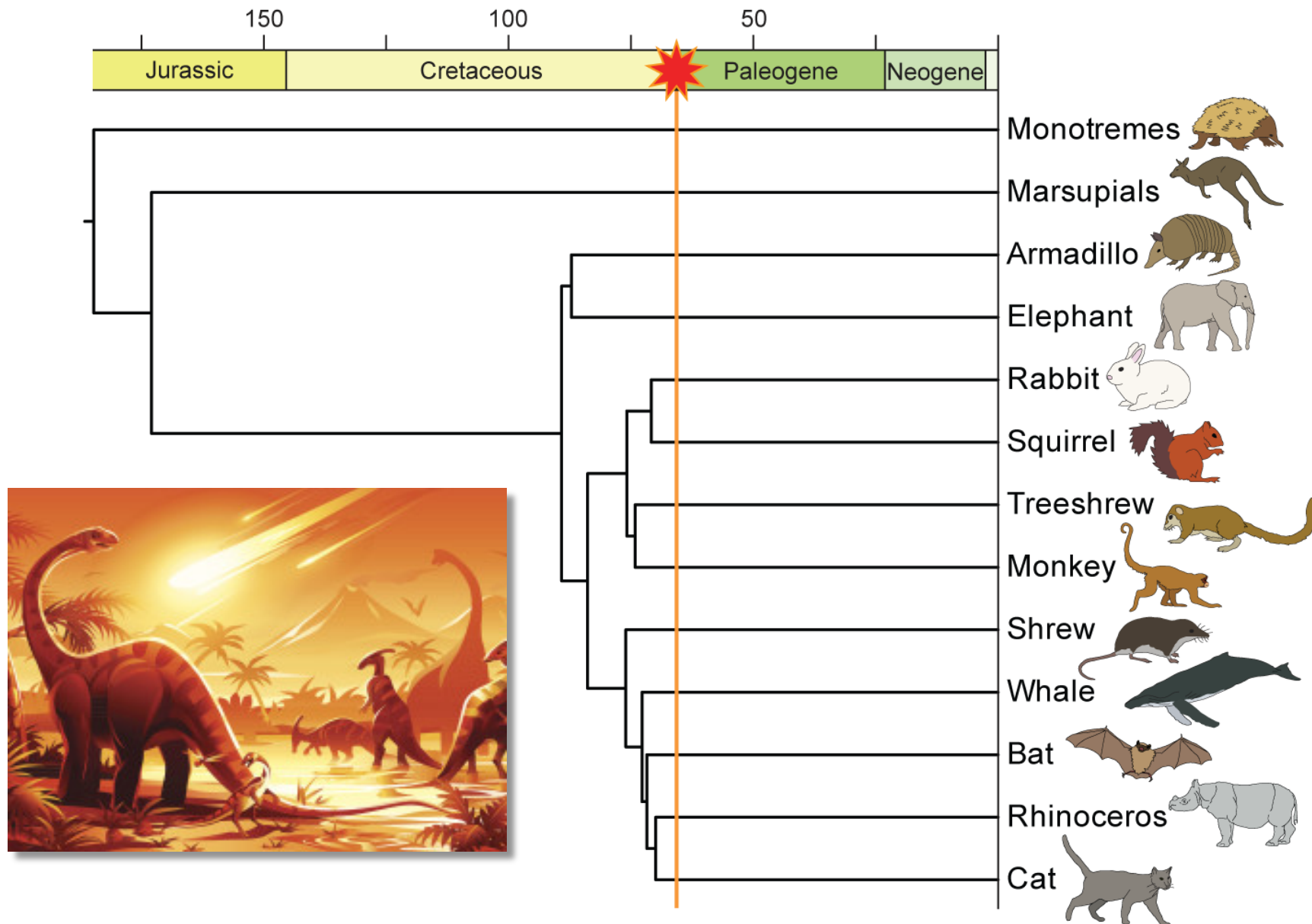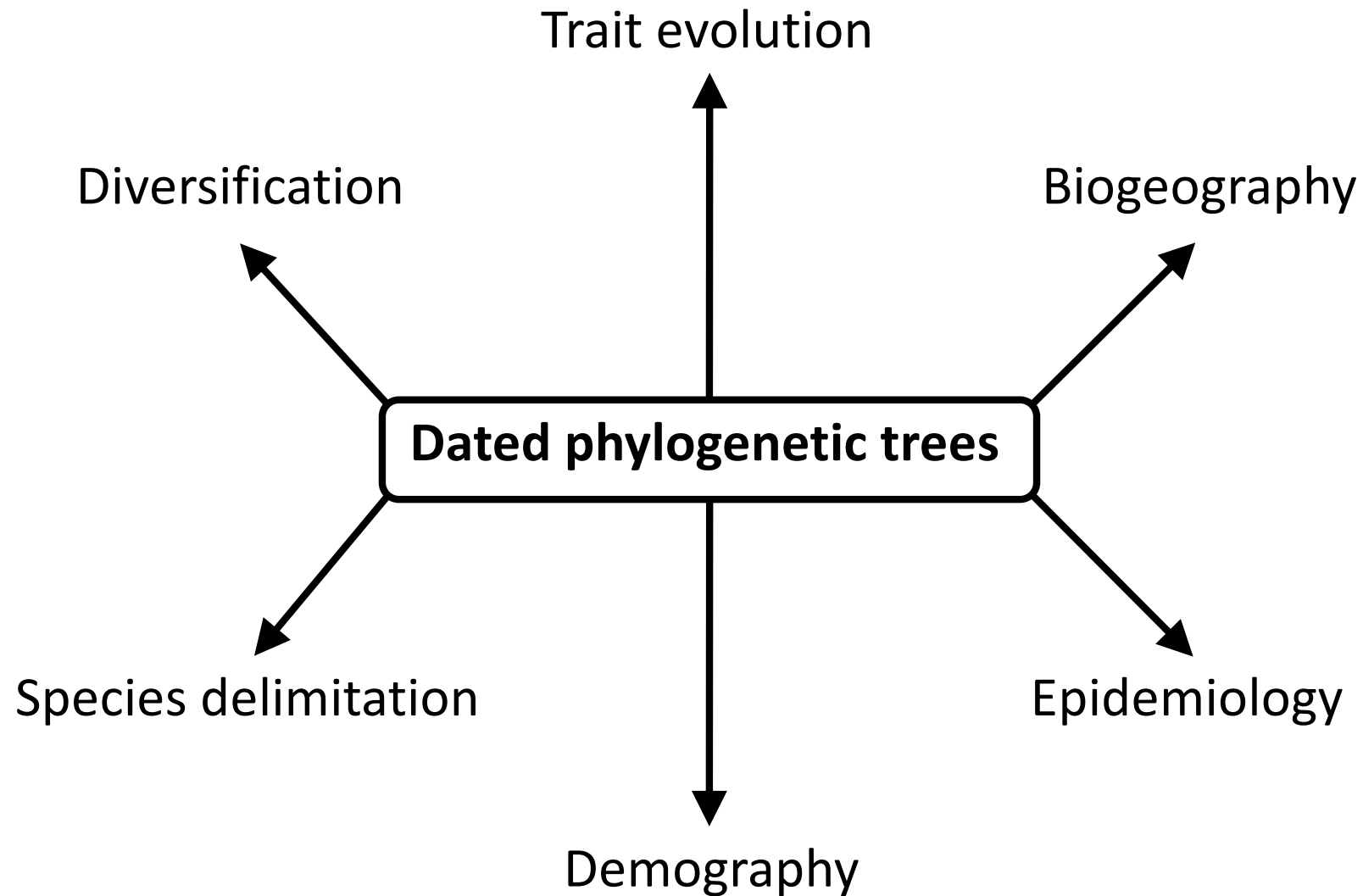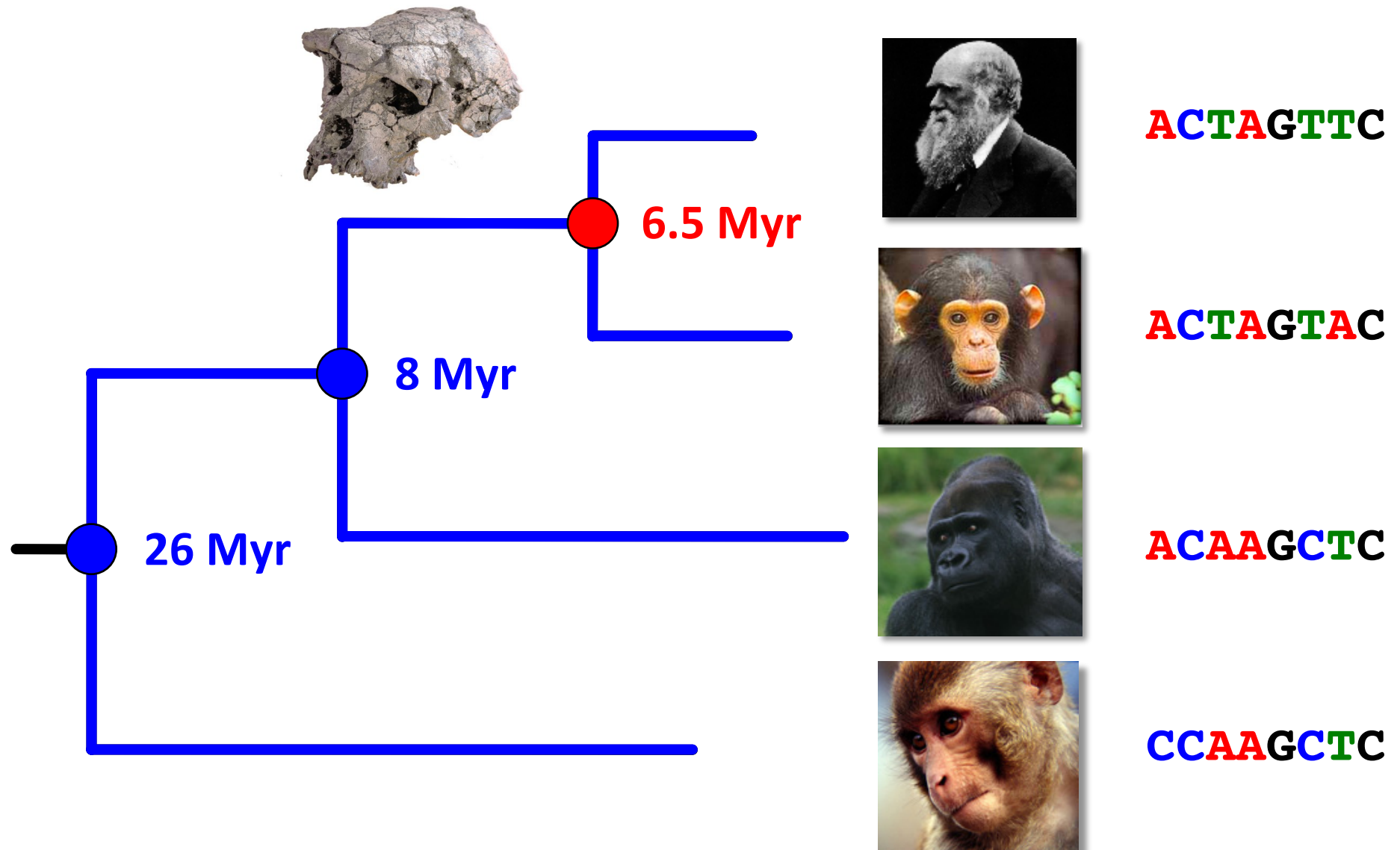

dos Reis *et al.* (2012) *Proc Roy Soc B*

Evolutionary timescales

Trait evolution

Diversification

Biogeography

Dated phylogenetic trees

Species delimitation

Epidemiology

Demography

The Molecular Clock

The molecular clock

The molecular clock

ACTAGTTC

ACTAGTAC

6.5 Myr

8 Myr

26 Myr

ACAAGCTC

CCAAGCTC

Sources of error

3. Calibration

1. Tree

6.5 Myr

8 Myr

26 Myr

ACTAGTTC

ACTAGTAC

ACAAGCTC

CCAAGCTC

2. Model of rate variation
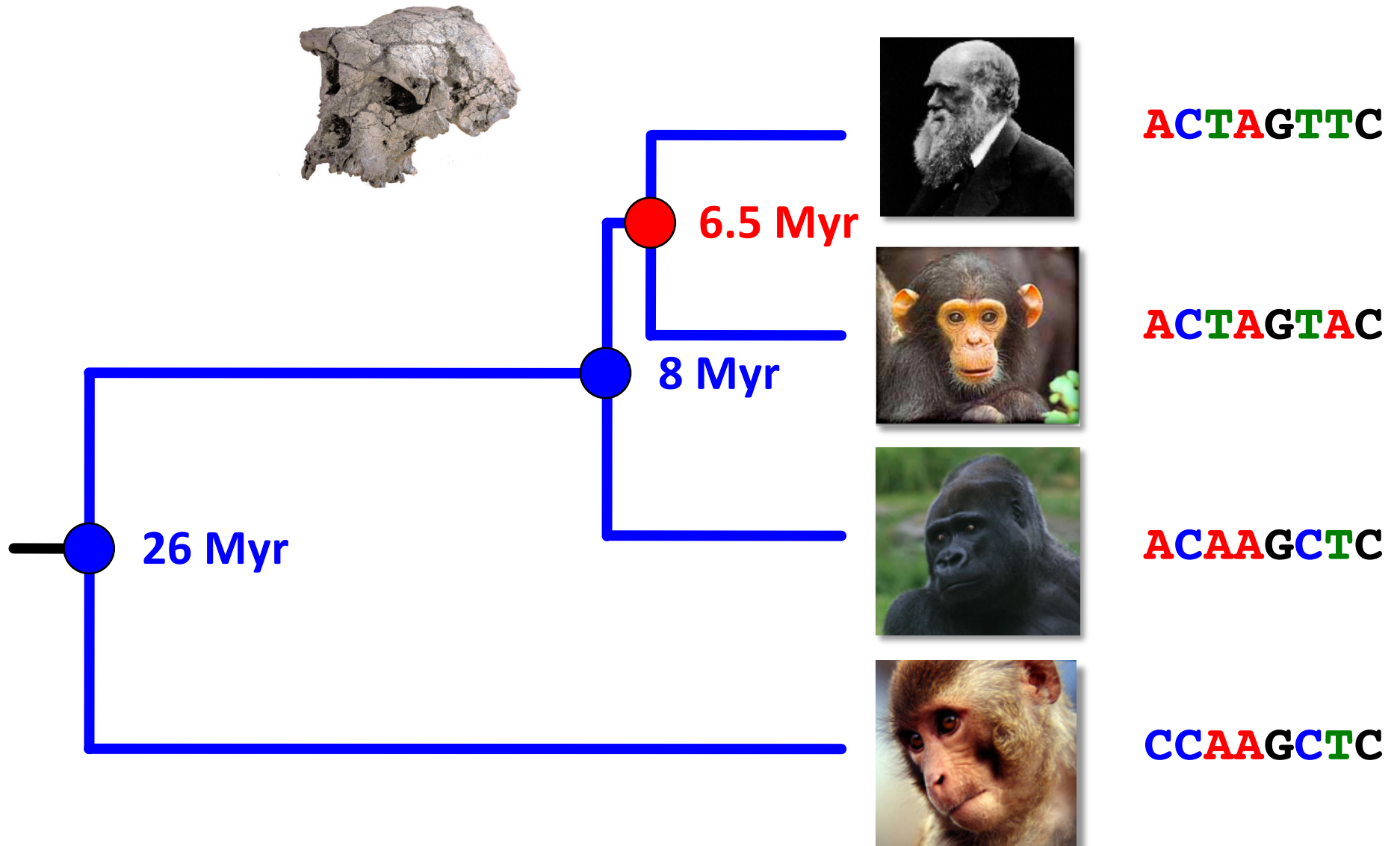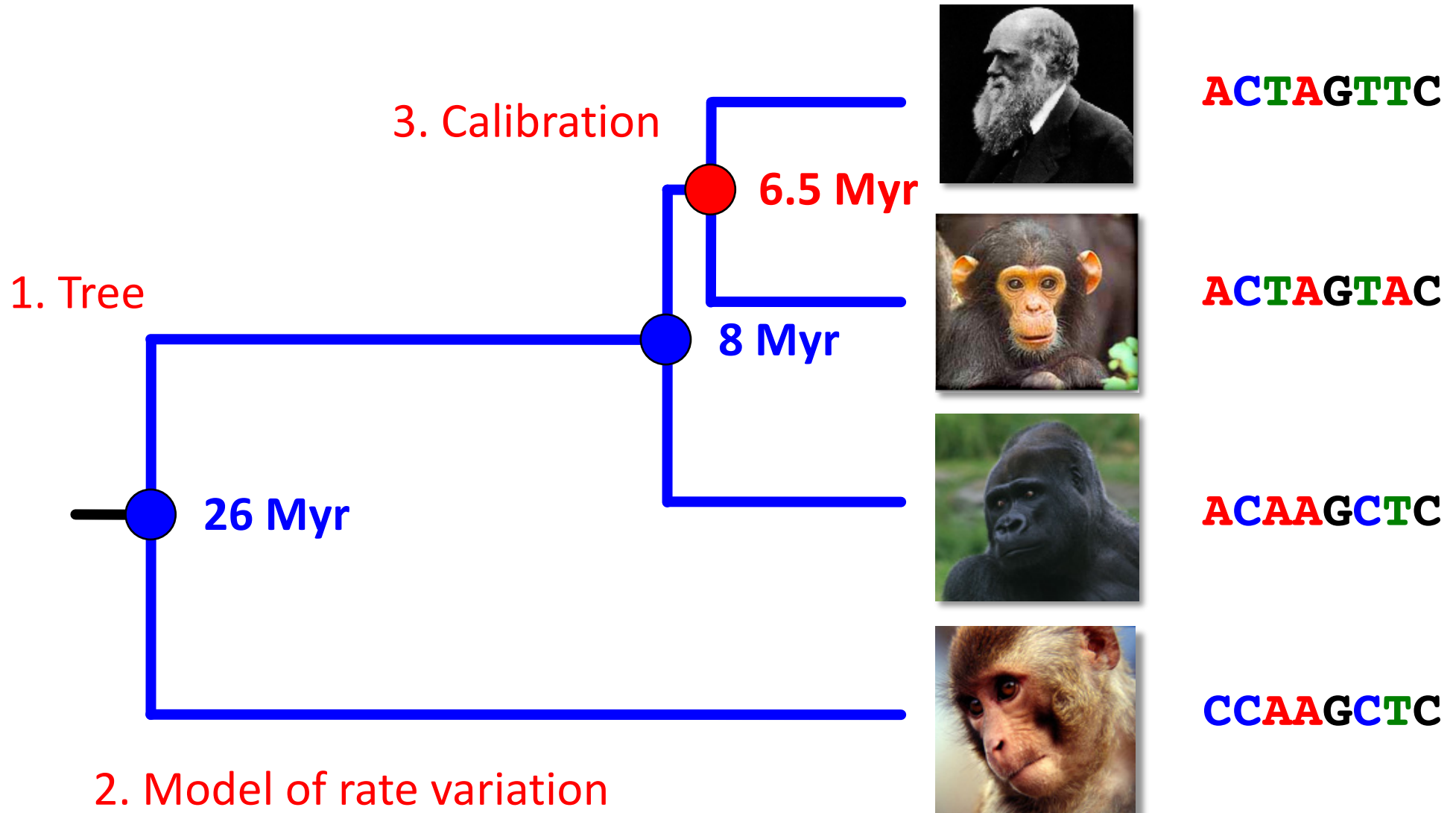
A brief history

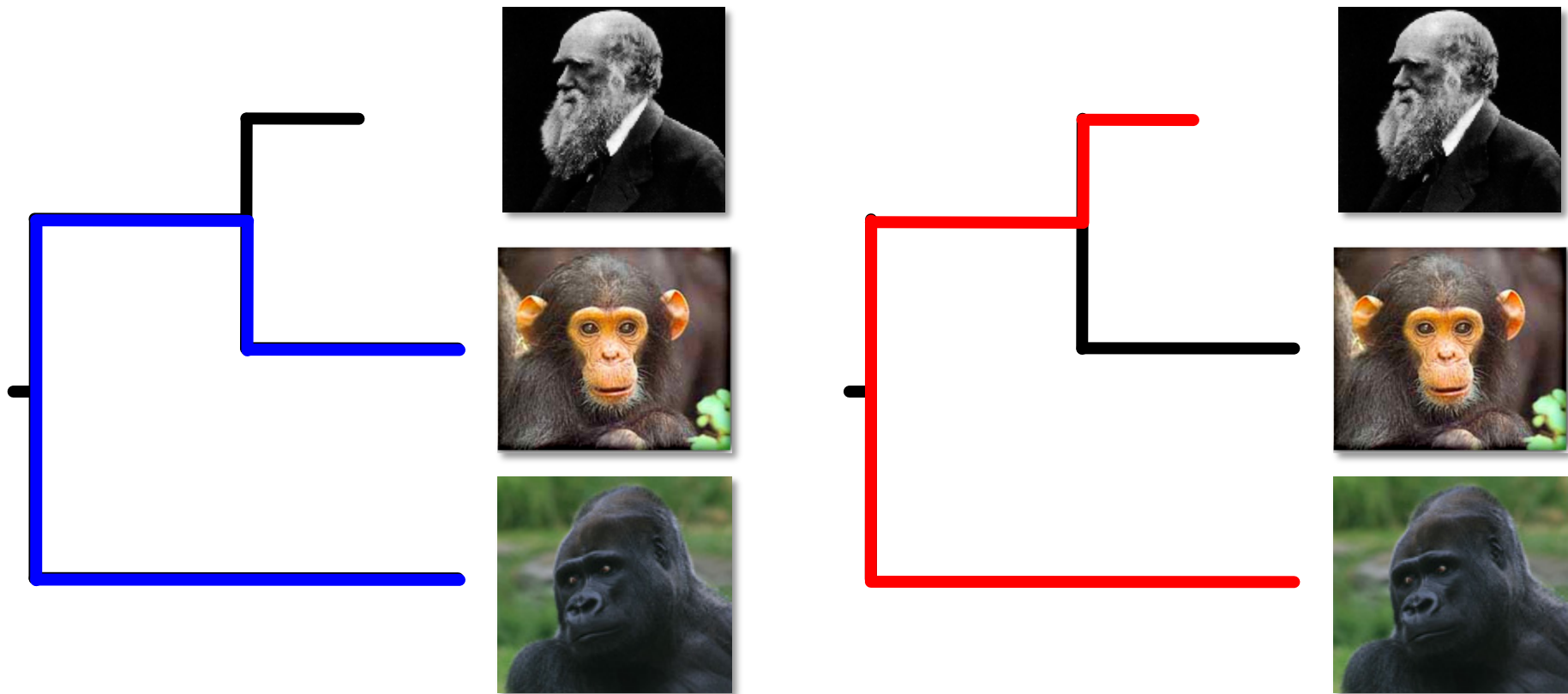- Emile Zuckerkandl and Linus Pauling

  1962   Assumed constant rate among species to estimate timing of globin gene duplications

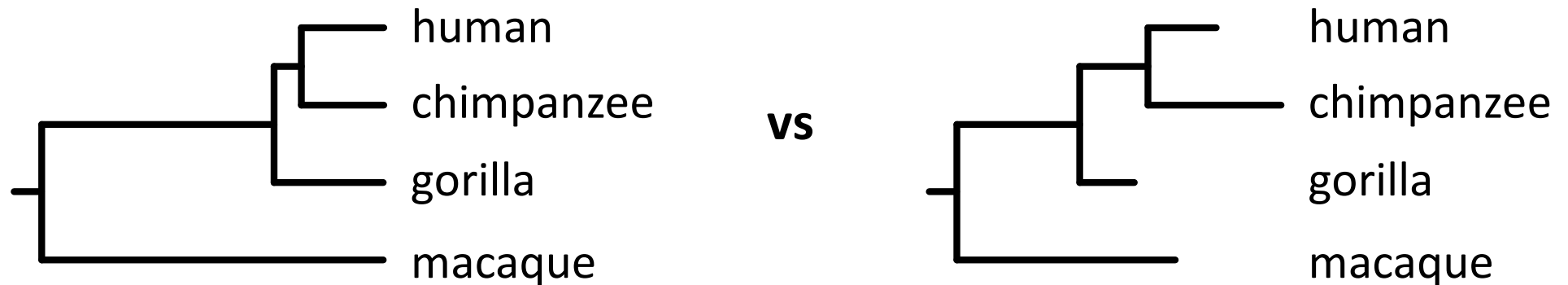 **1965**   Introduced the term 'molecular evolutionary clock'

Testing for clocklike evolution

- Relative-rates test (Fitch, 1976)

Testing for clocklike evolution

- Whole-tree methods (*e.g.*, likelihood-ratio test)

- Limitations
  - Cannot identify instances in which all lineages experience simultaneous rate shift
  - Failure to reject clock can be due to lack of power or lack of information

Departures from the clock

- Rates vary among lineages because of differences in:

  - Exposure to mutagens

  - Metabolic rate

  - Generation time

  - Population size
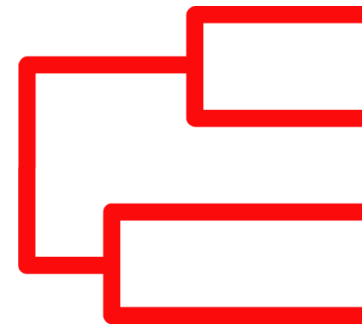
  - Strength and direction of selection
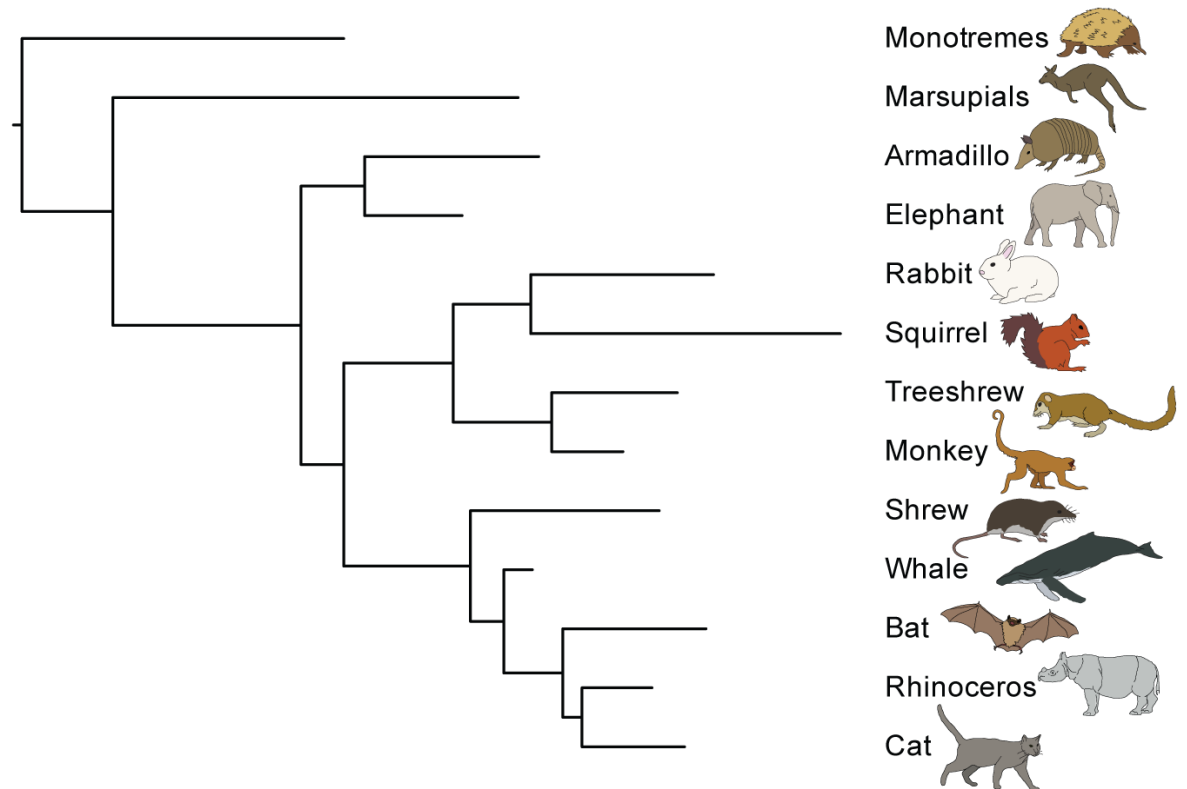
The assumption of rate constancy can be relaxed
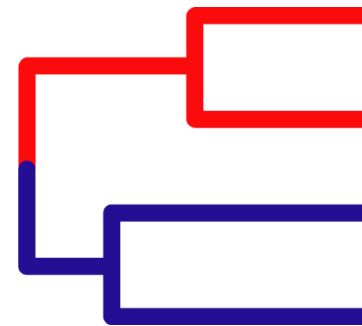
Molecular-Clock Models
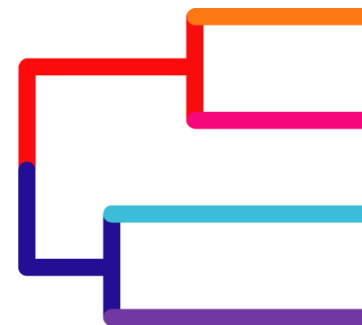
Molecular-clock models

Strict or 'global' molecular clock

$k = 1$

Multi-rate clocks

$1 < k < n$

Relaxed clocks

$k = n$

Multi-rate clocks

- Small number of rates

  - More than 1 rate (*i.e.*, not a strict clock)

  - Fewer than number of branches (*i.e.*, not a relaxed clock)

- **Local clock**

  - Same rate shared by neighbouring branches

- **Discrete clock**

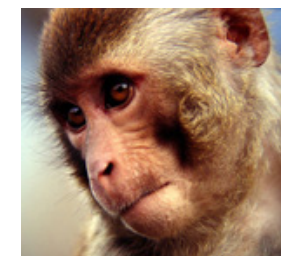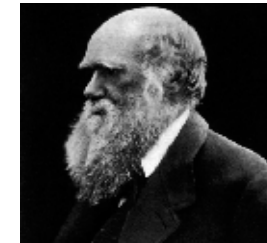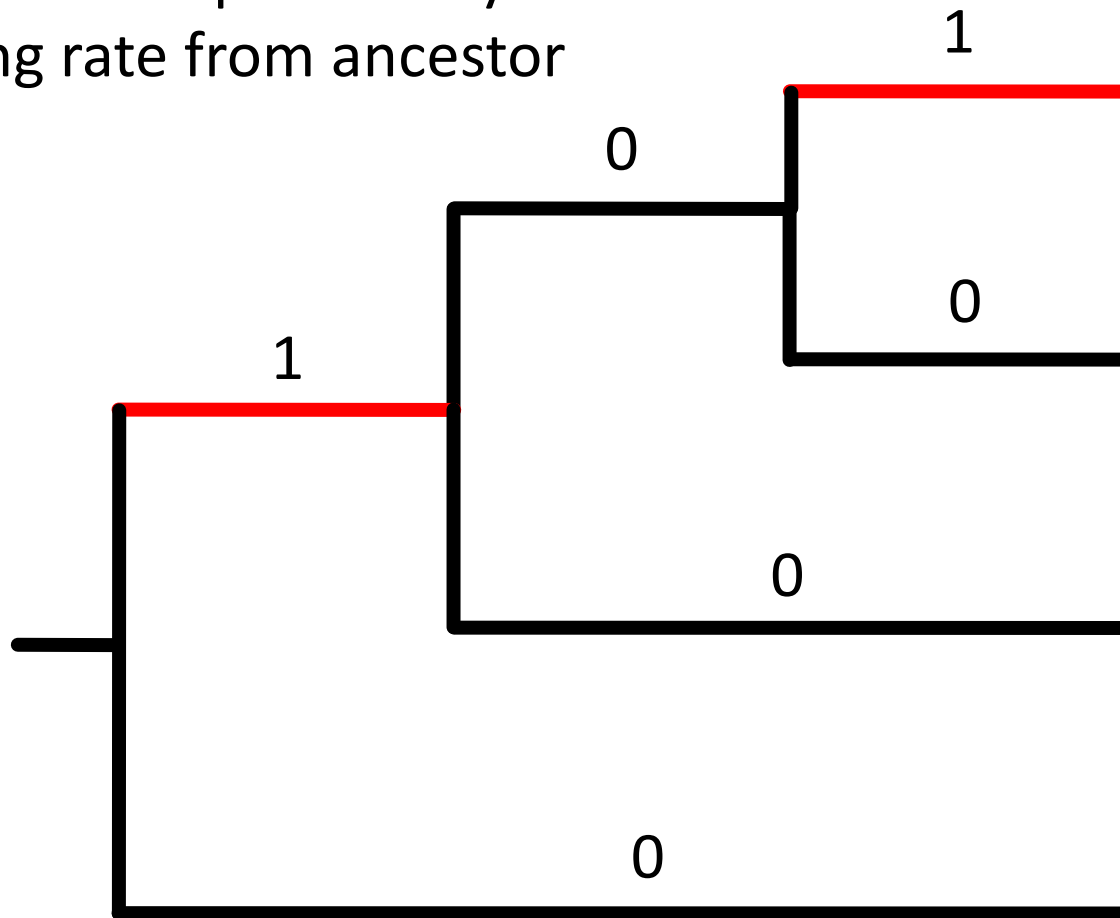 - Small number of branch rates, distributed across tree

Local clocks

- **User-defined local clock**

  - Fixed tree topology

- **Random local clock**

  - Each branch has a probability of inheriting rate from ancestor
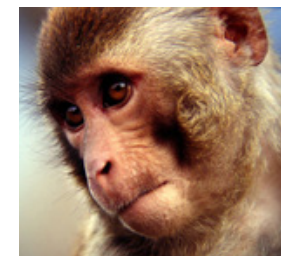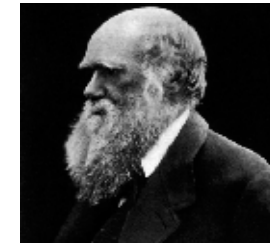
  - Tree estimated
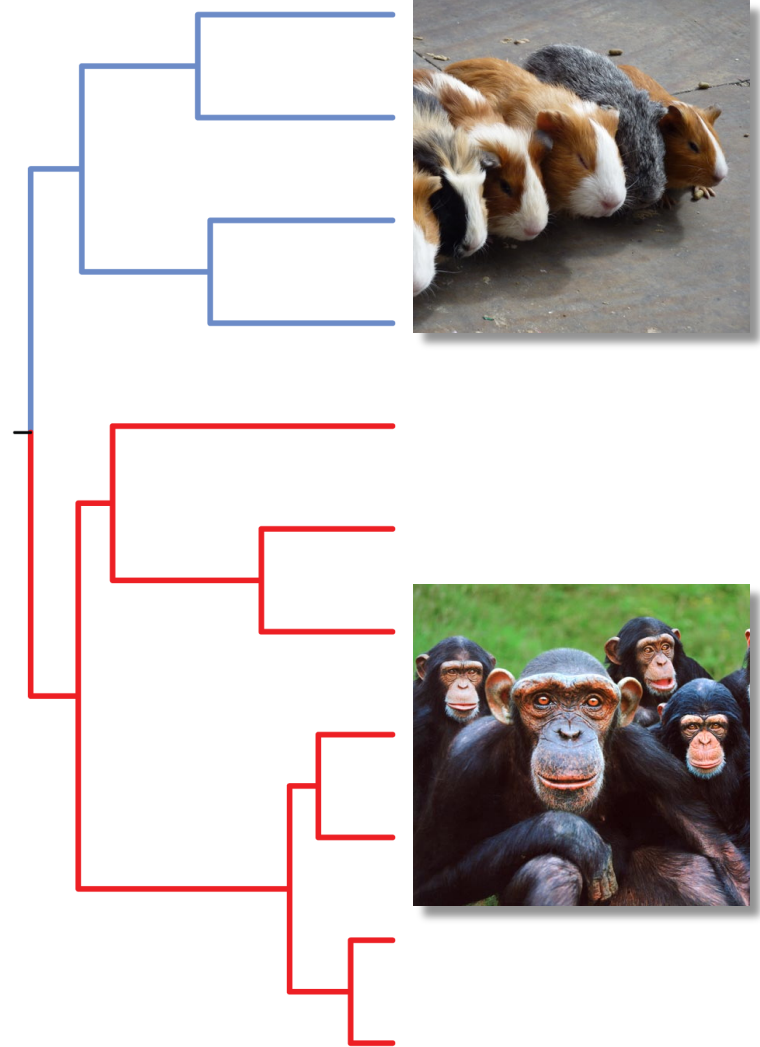
Random local clock

Each branch has a probability
of inheriting rate from ancestor

Includes possibility of zero rate changes (= strict clock)

Random local clock

Otherwise multiplied by a relative rate modifier

μ_2

μ_1

μ_1

μ_1

μ_1

μ_0
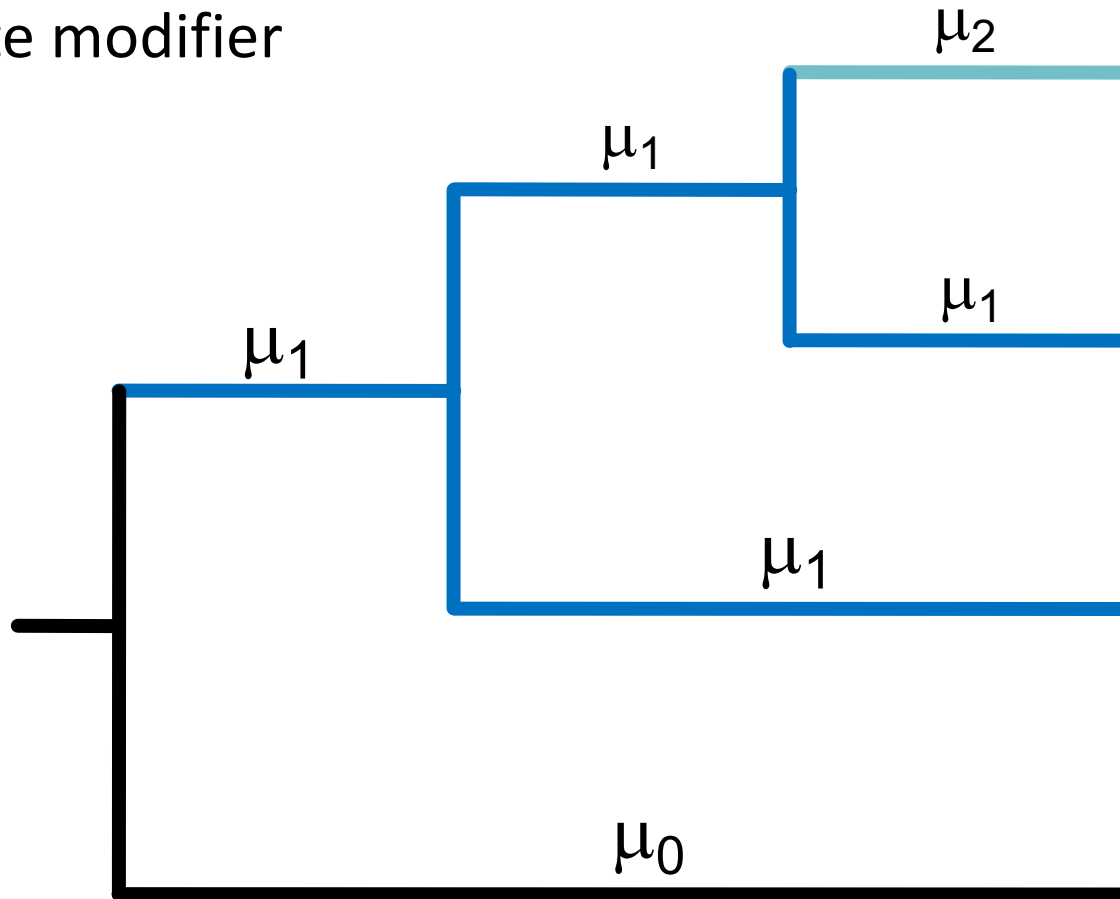
Note that tree is jointly estimated

Relaxed clocks

- We know that life-history characteristics:

  - Have effects on rates of molecular evolution

  - Are usually heritable to some degree

- Treat molecular rate as a heritable trait

- Relaxed clocks generally assume that closely related species share similar rates

Likelihood-based relaxed clocks

- Allow a different rate in each branch

- Penalise large, sudden changes in rates

Large rate change
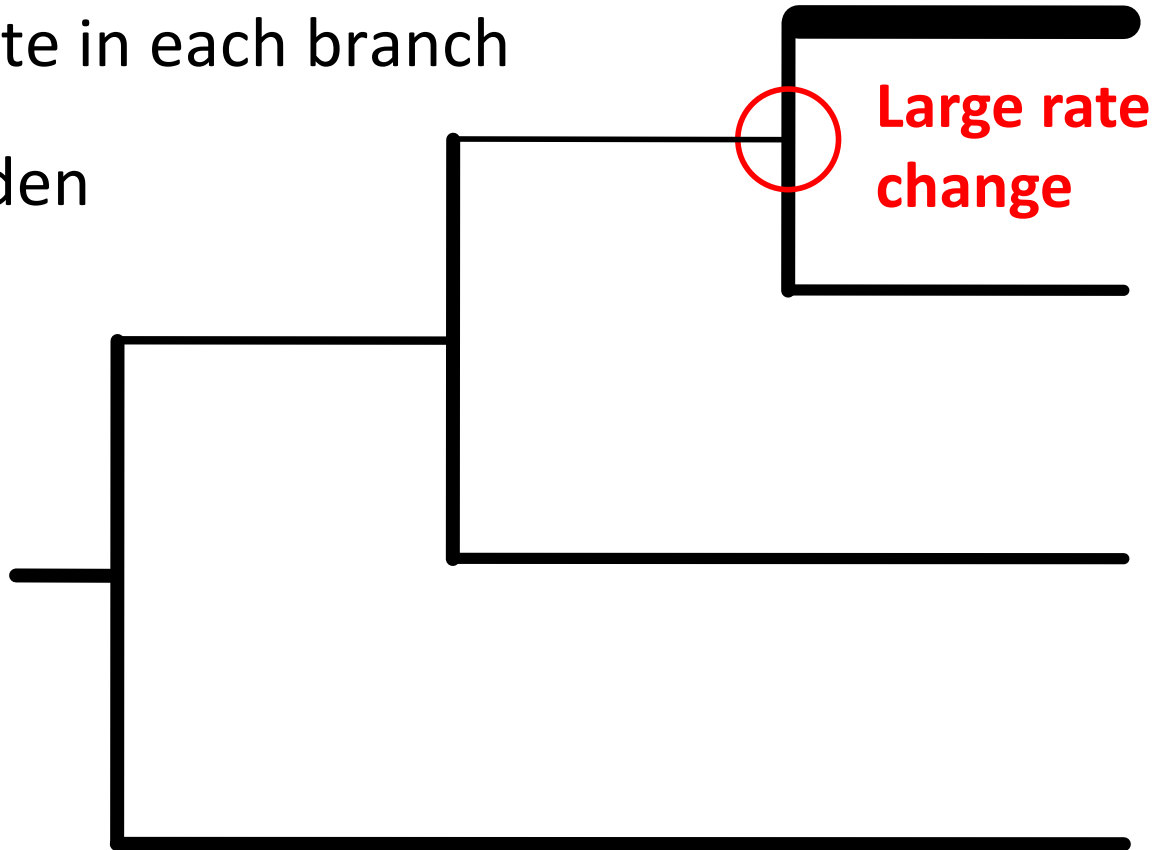
Penalised likelihood = lnL − $\lambda\Phi(r)$

Bayesian relaxed clocks

- Allow a different rate in each branch

- Statistical models of rates among branches

- Rates can be autocorrelated or uncorrelated

  - **Autocorrelated**
    rates in neighbouring branches are related

  - **Uncorrelated**
    rates identically and independently distributed among branches

Bayesian relaxed clocks

Rate *j*

Rate *i*

Autocorrelated lognormal relaxed clock

Bayesian relaxed clocks

Rate *j*

Rate *i*
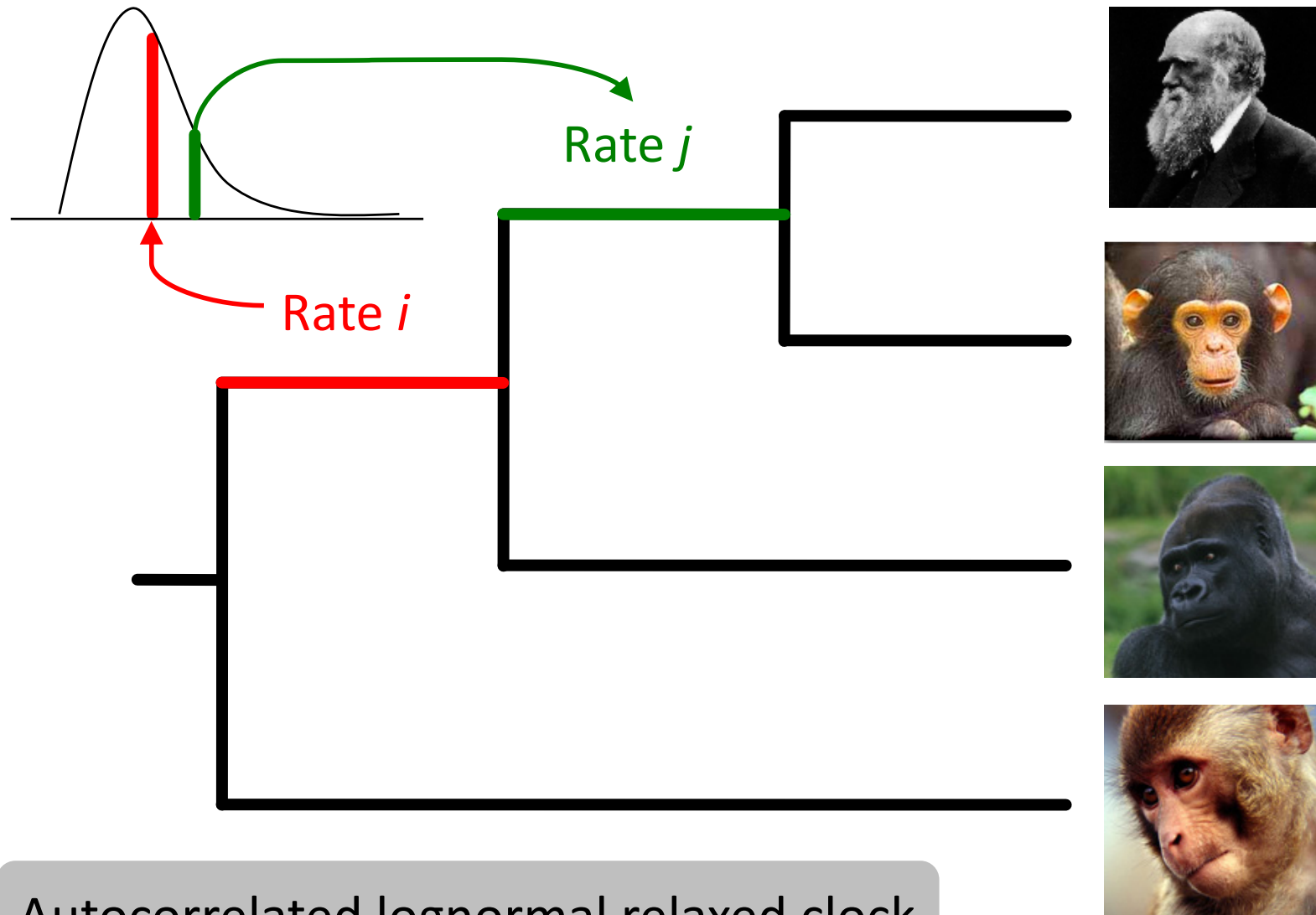
Uncorrelated lognormal relaxed clock
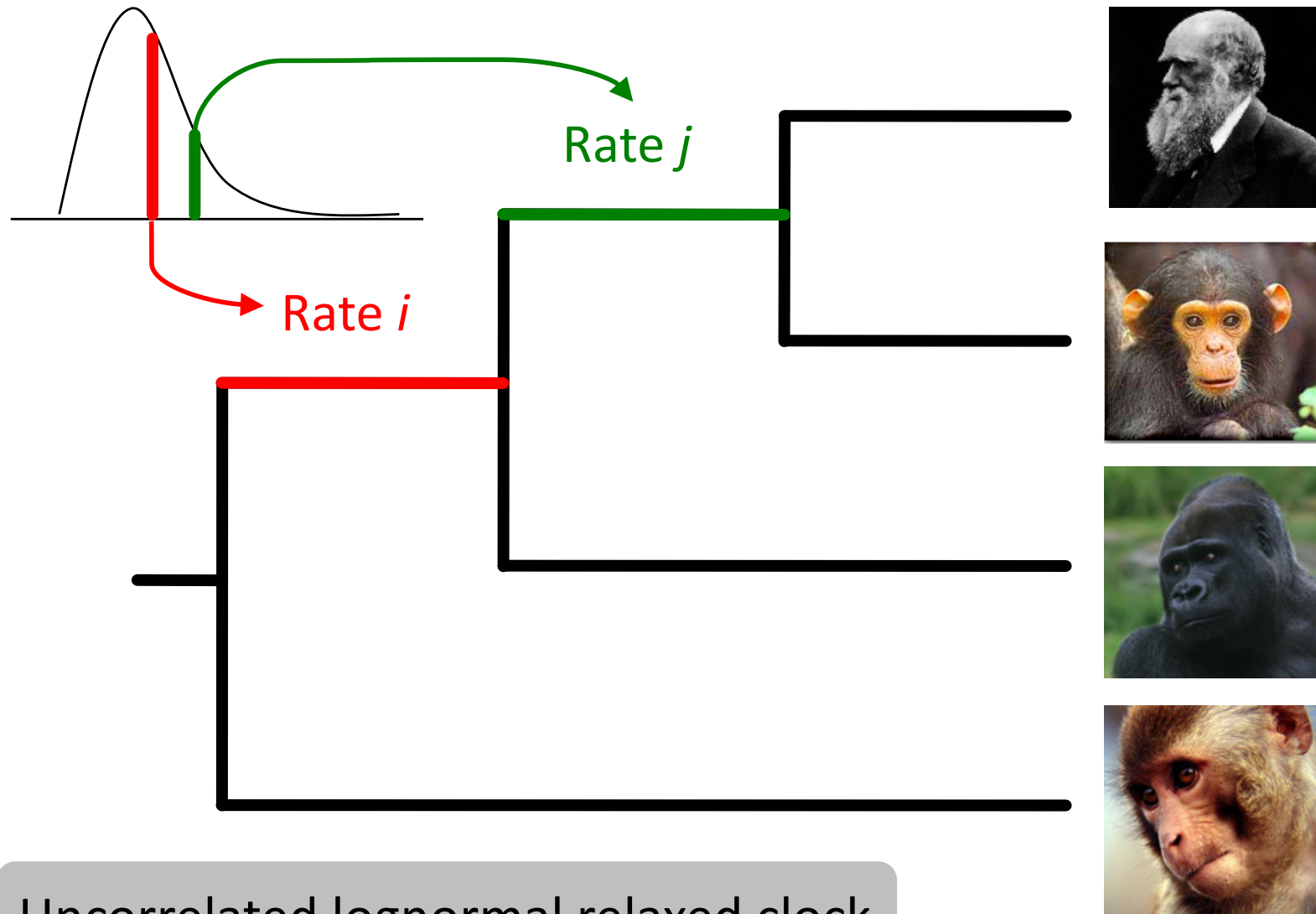
Bayesian relaxed clocks

- In the uncorrelated lognormal relaxed clock,
  two statistics can be obtained:

  1. **Coefficient of variation of rates**
     Measures the rate variation among branches
     A value of 0 indicates clocklike evolution

  2. **Covariance of rates**
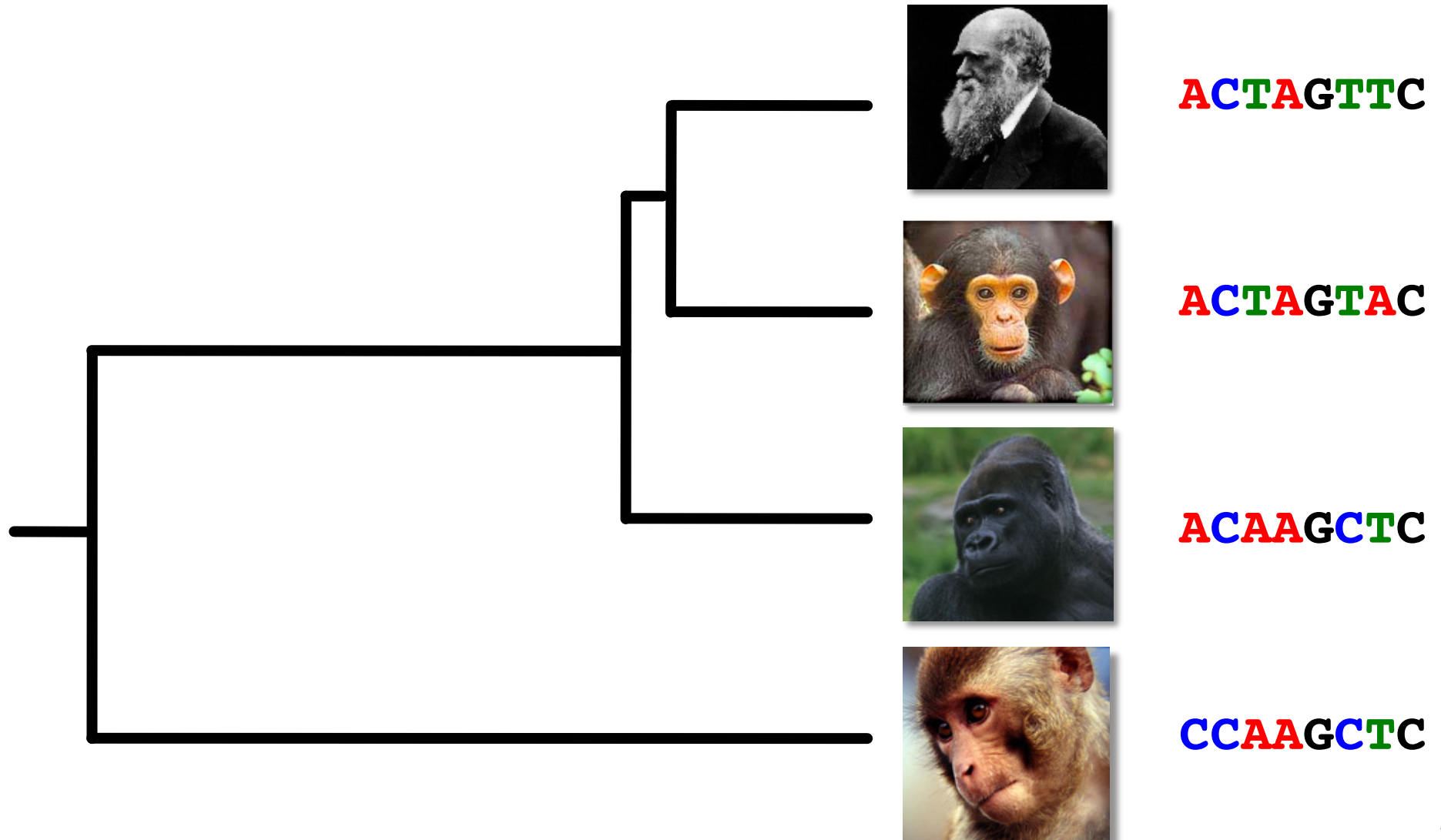   Measures autocorrelation of rates between adjacent branches

Rate autocorrelation

- Little evidence of rate autocorrelation in real data

- Uncorrelated relaxed clock probably appropriate in most cases

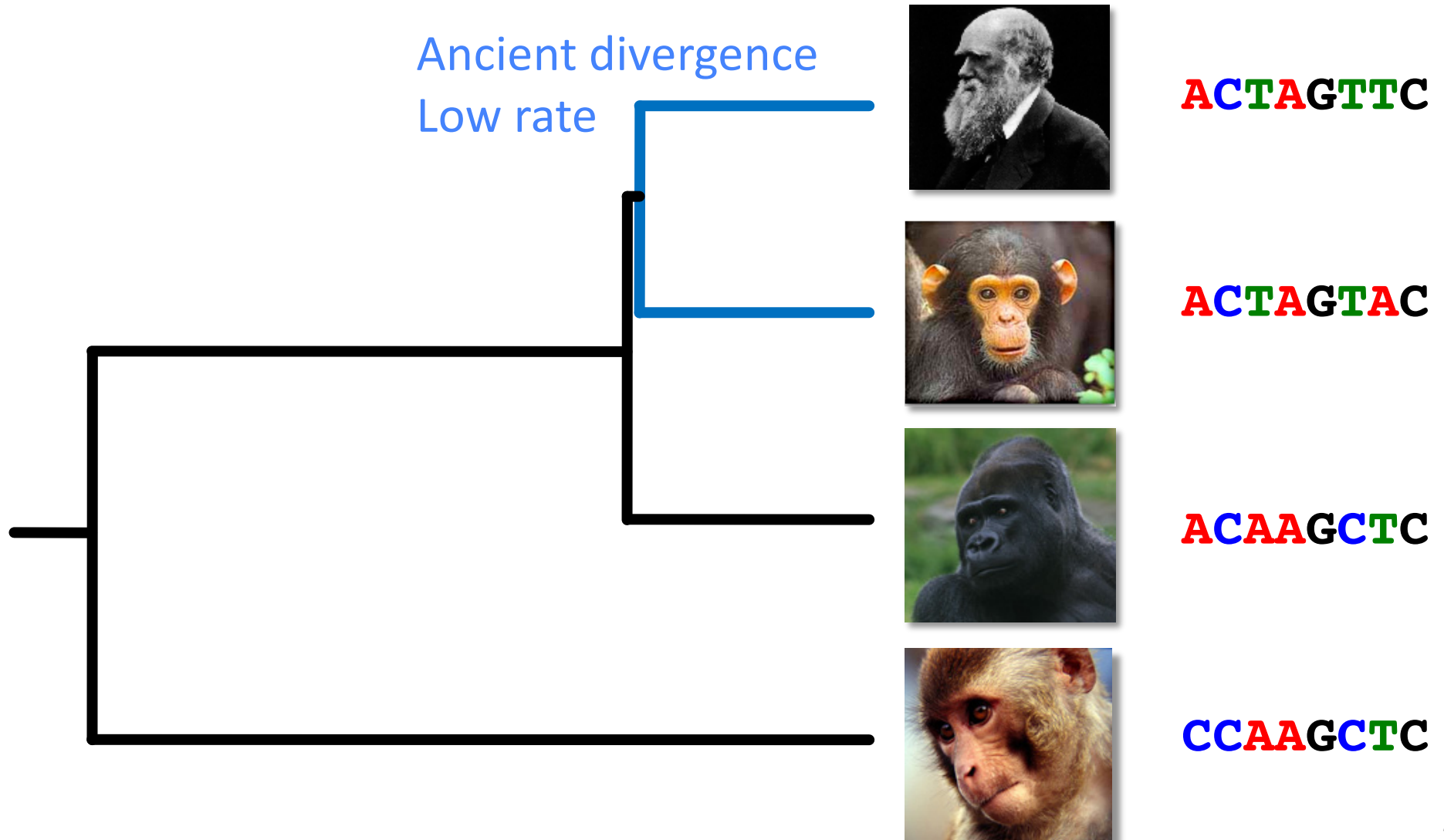- Compare estimates from auto- and uncorrelated relaxed clocks
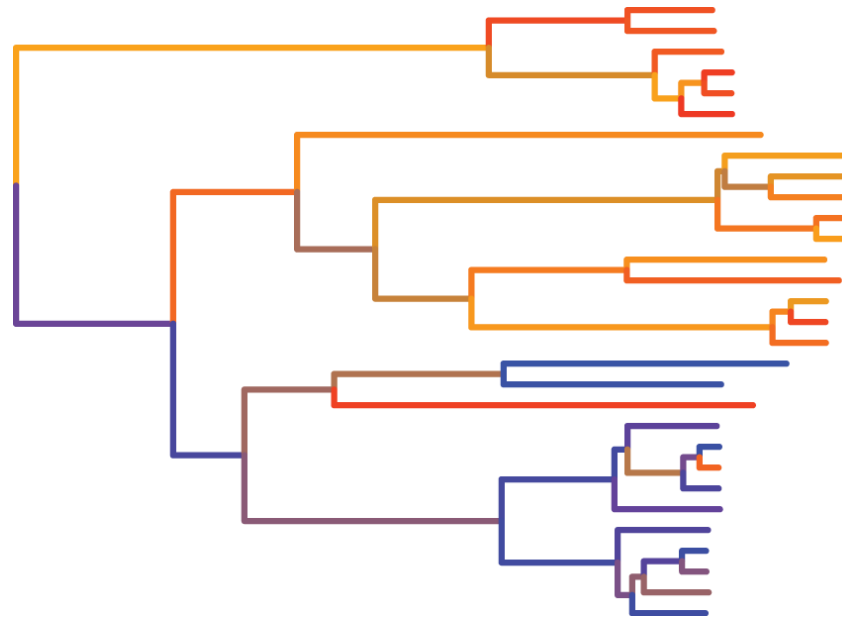
Ho, Duchêne, & Duchêne (2015) *Mol Ecol Resour* 24

Calibrating the Molecular Clock

Calibrating the molecular clock

ACTAGTTC

ACTAGTAC

ACAAGCTC

CCAAGCTC

Calibrating the molecular clock

Calibrating the molecular clock

Recent divergence
High rate

ACTAGTTC

ACTAGTAC

ACAAGCTC

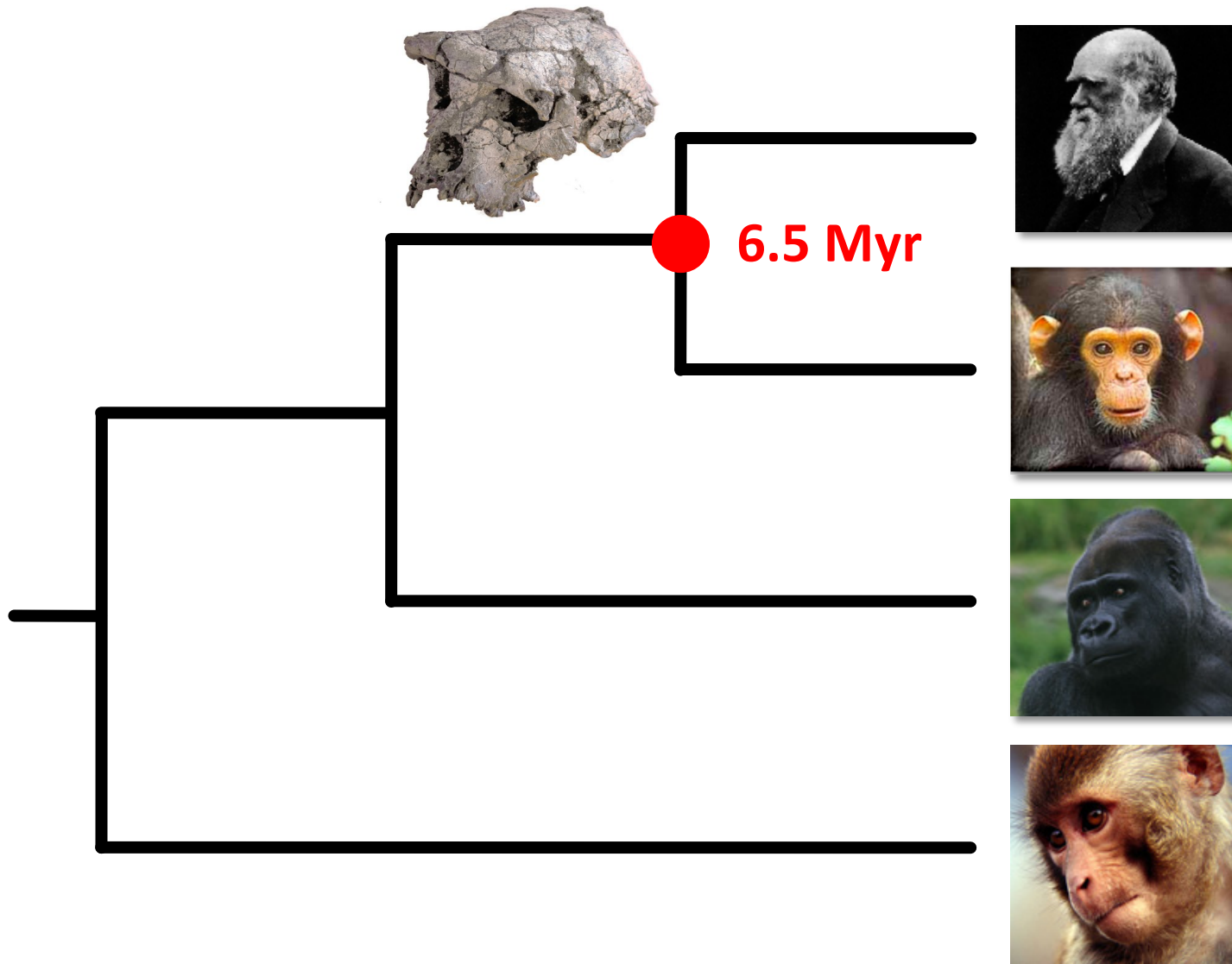**CCAAGCTC**
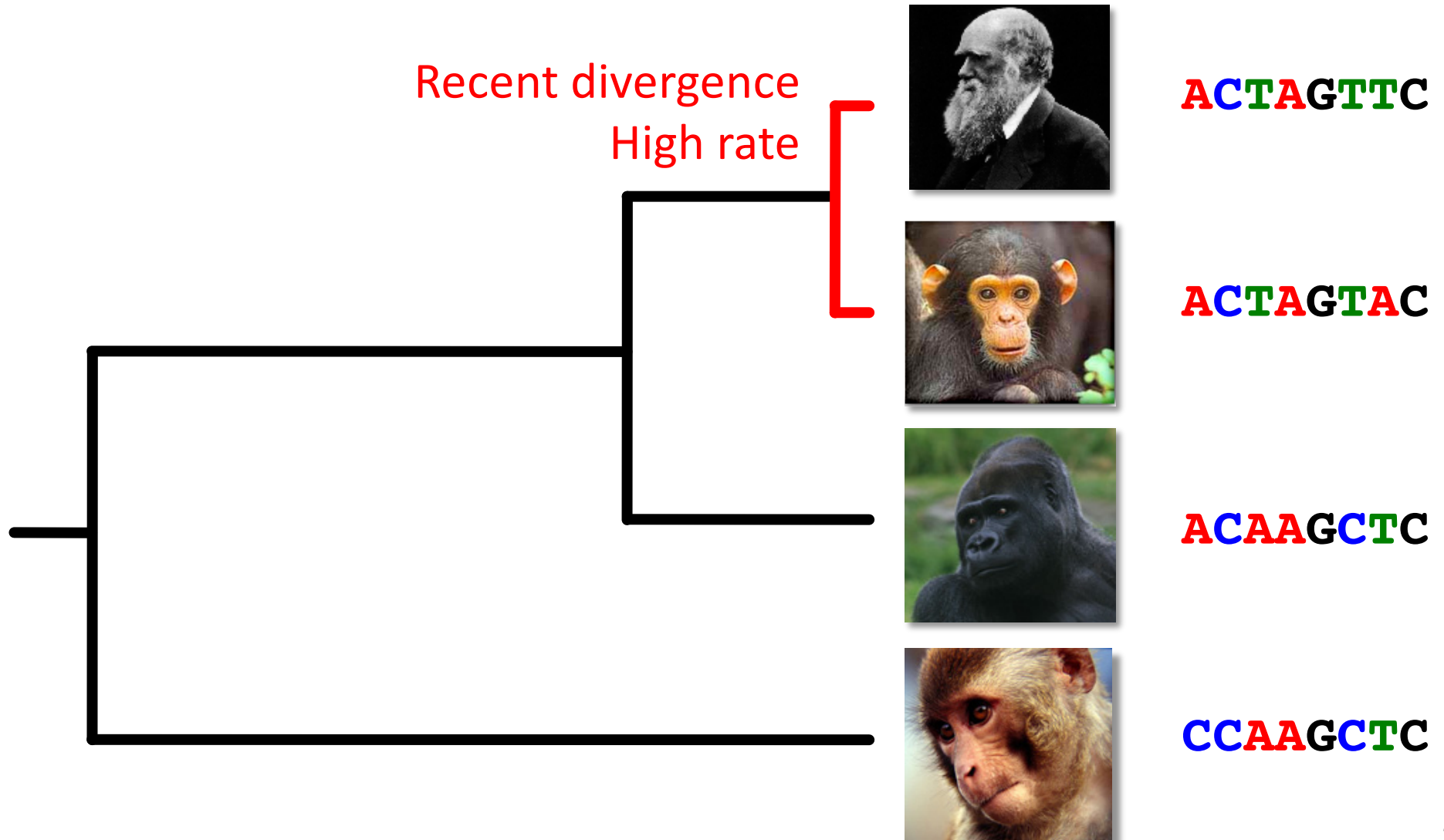
Calibrating the molecular clock

- Rates and times are **non-identifiable**

- Likelihood only depends on their product
  - Branch lengths in substitutions per site

- To separate rate and time, we need (prior) information about one or the other

Calibration: Fossil record

6.5 Myr

Useful references

- **A practical guide to molecular dating**
  Sauquet (2013) *Comptes Rendus Palevol,*
  12: 355–367.

- **Bayesian molecular clock dating of species divergences in the genomics era**
  dos Reis, Donoghue, & Yang (2016) *Nature Reviews Genetics,* 17: 71–80.

- **Molecular-clock methods for estimating evolutionary rates and timescales**
  Ho & Duchêne (2014) *Molecular Ecology,*
  23: 5947–5965.