Divergence time estimation
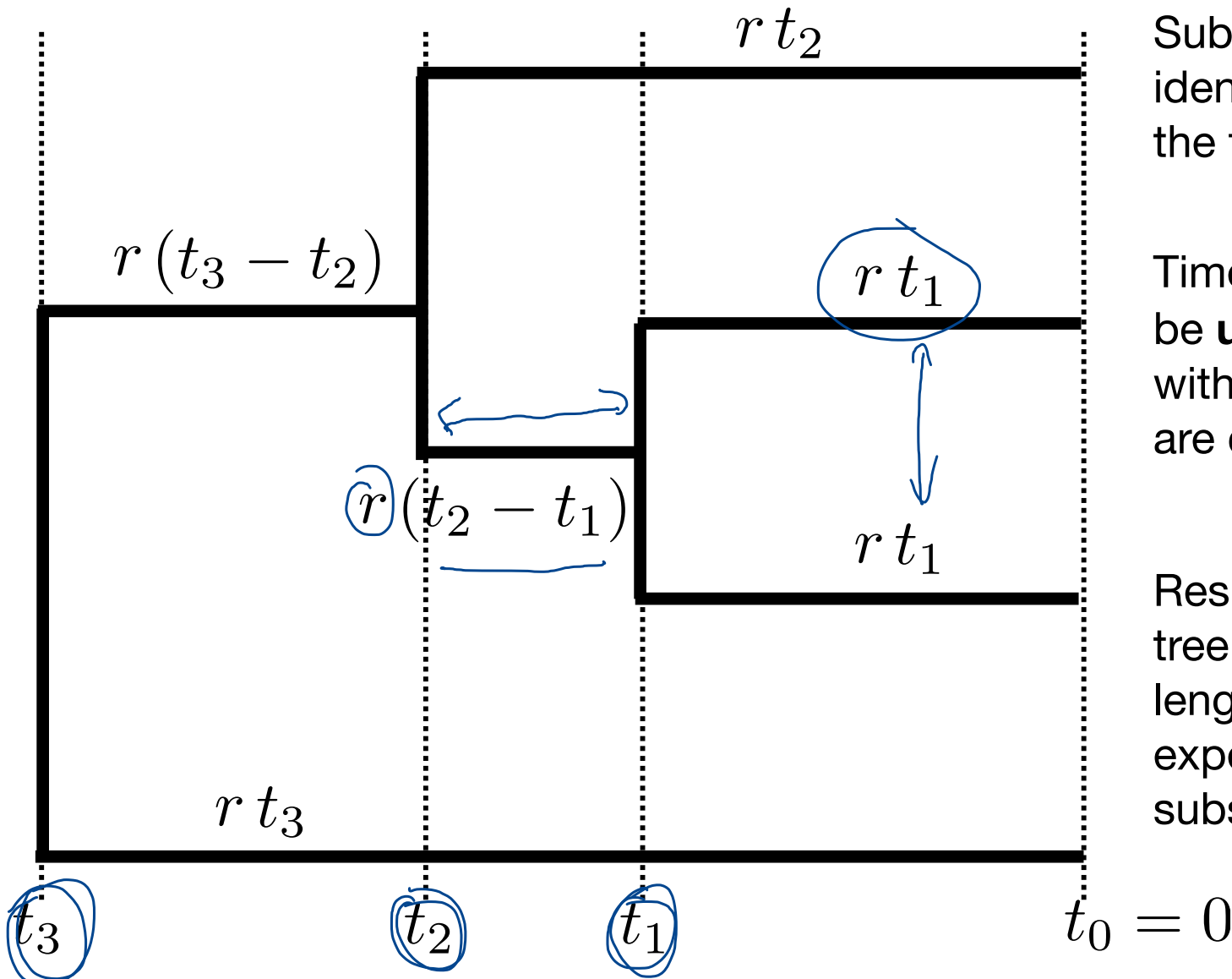
rate × time = expected number of substitutions

Strict Molecular Clock

r

r

r

r

r

r
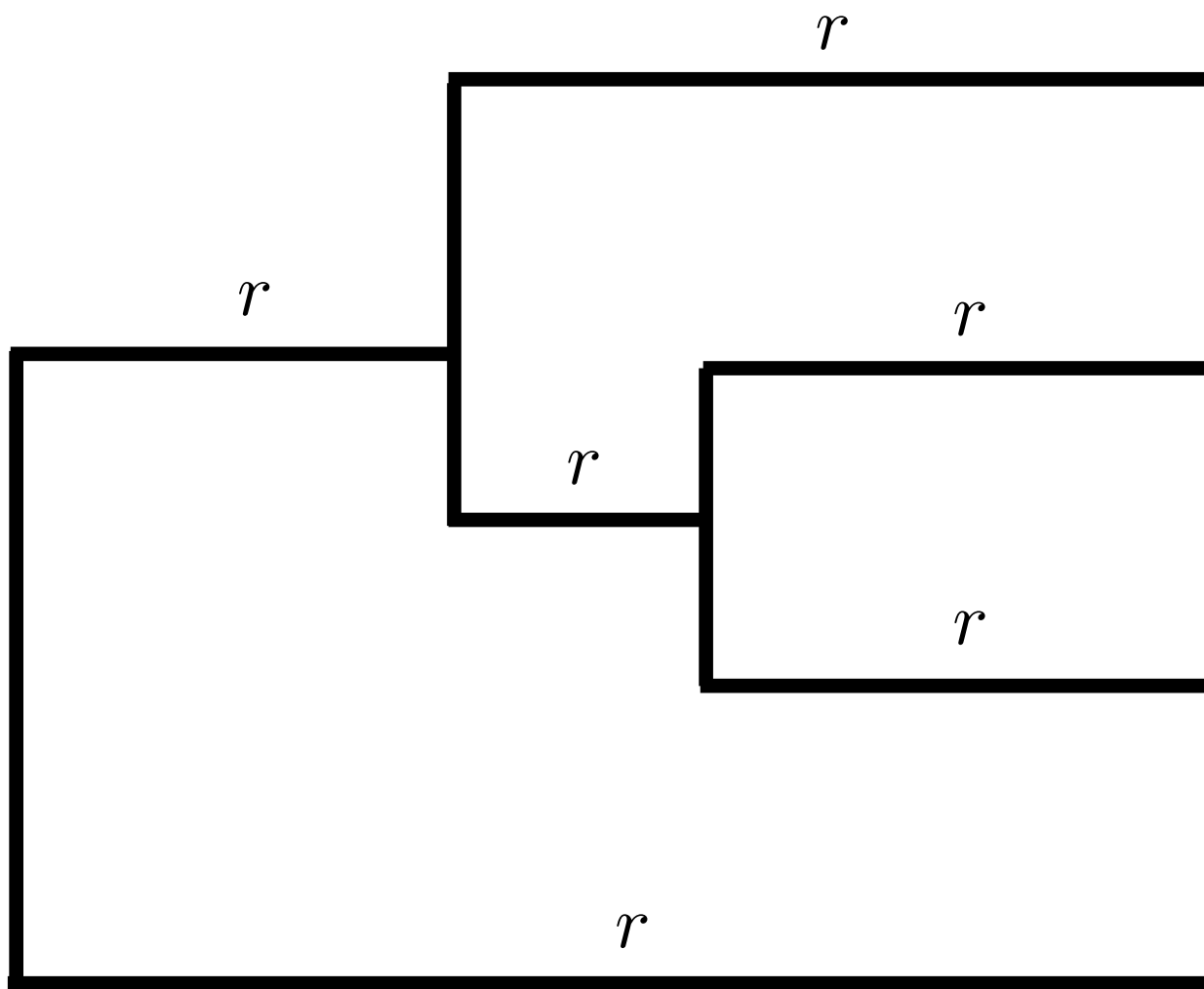
Substitution rate r is identical for every edge in the tree.

Strict Molecular Clock

$$r\,t_2$$

$$r\,(t_3 - t_2)$$

$$r\,t_1$$

$$r\,(t_2 - t_1)$$

$$r\,t_1$$

$$r\,t_3$$

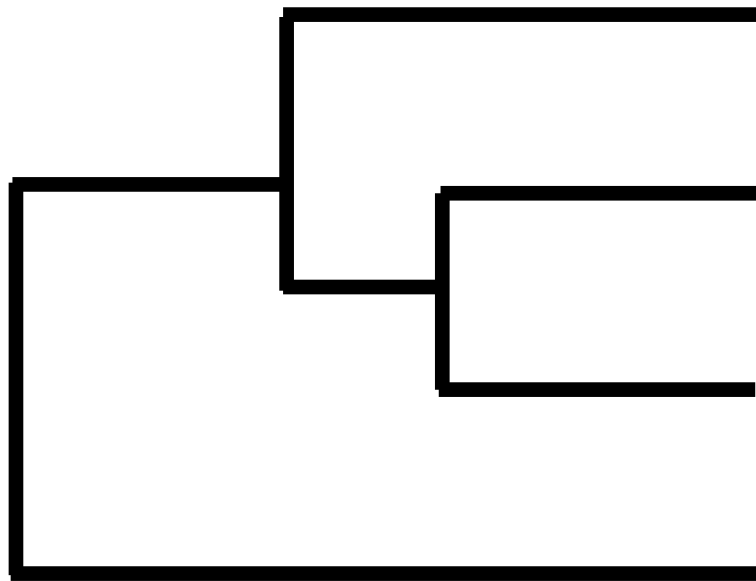$$t_3 \qquad t_2 \qquad t_1 \qquad t_0 = 0$$

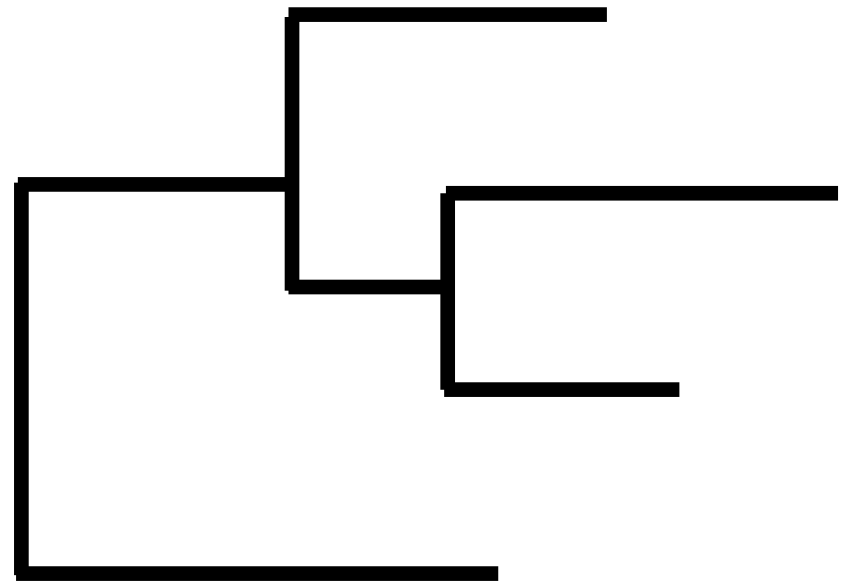Substitution rate *r* is identical for every edge in the tree.

Times are constrained to be **ultrametric**, consistent with the fact that all tips are contemporary

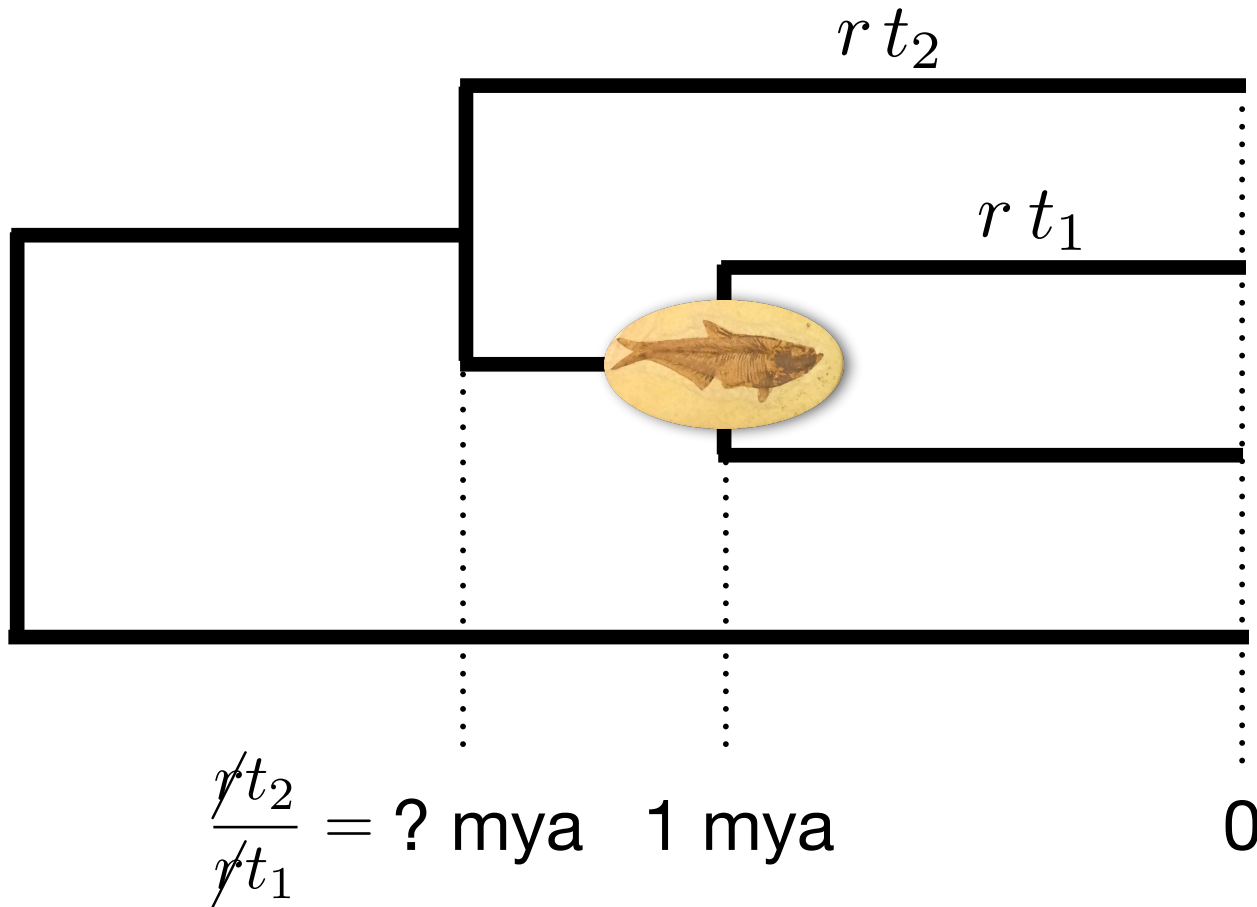Result is an ultrametric tree even when edge lengths are measured in expected number of substitutions per site (*rt*)
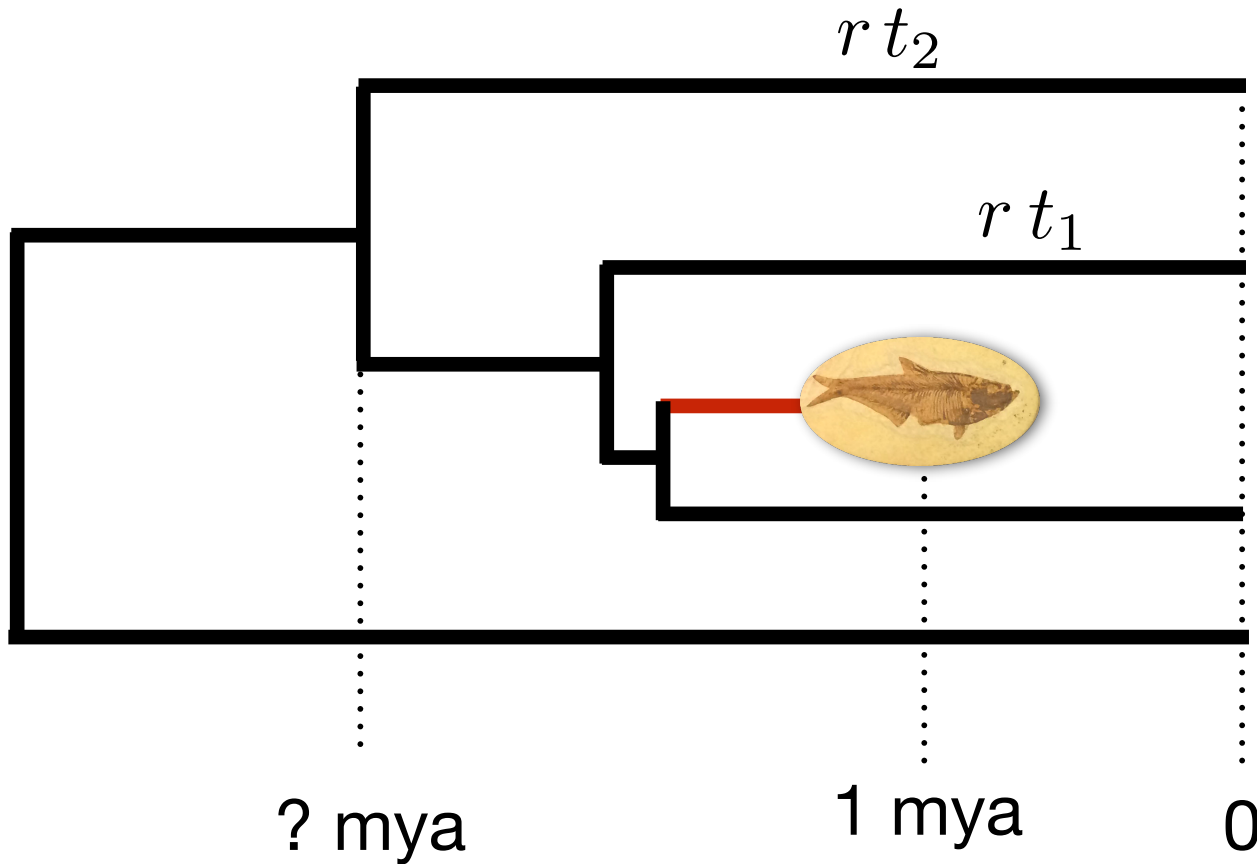
strict clock

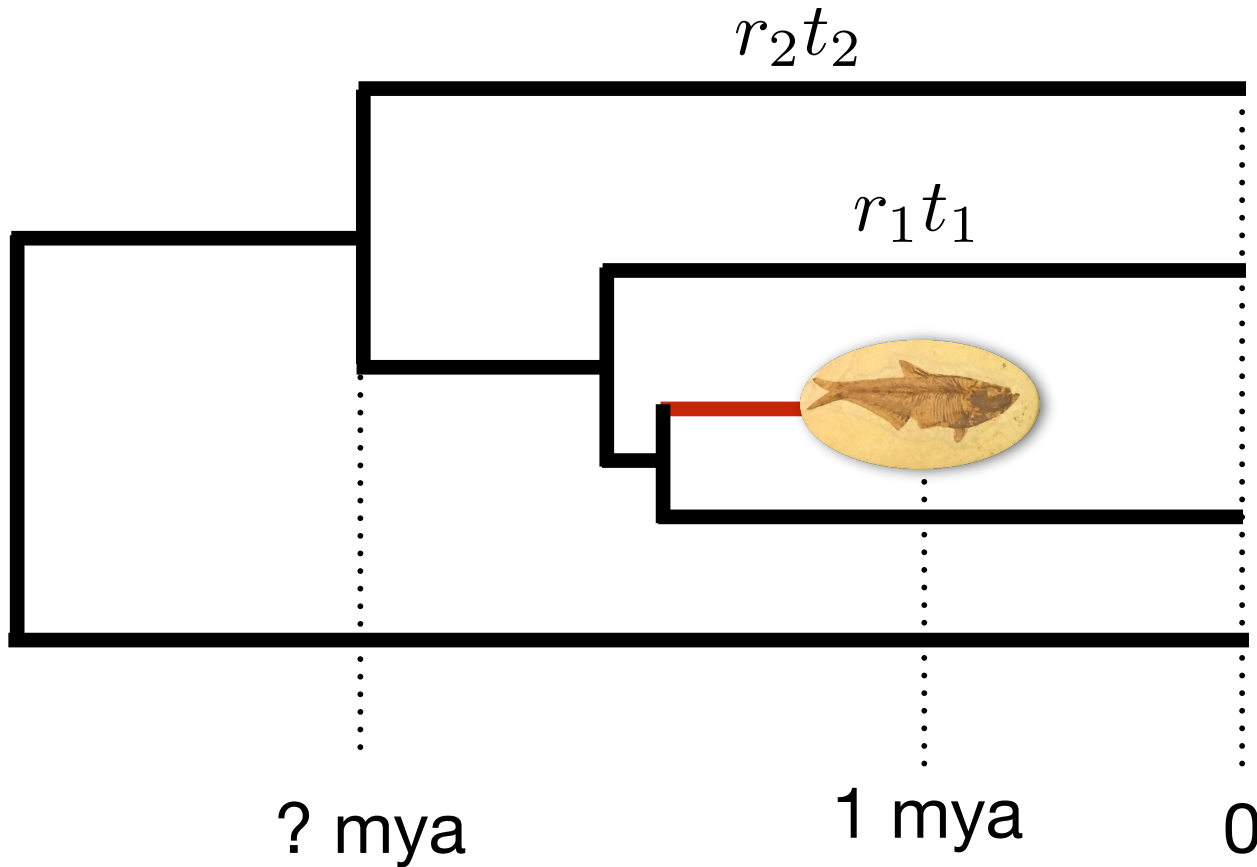# relaxed clock

Calibrated Strict Clock

$$r\,t_2$$

$$r\,t_1$$

$$\frac{\cancel{r}\,t_2}{\cancel{r}\,t_1} = ?\ \text{mya} \quad 1\ \text{mya} \qquad\qquad 0$$

Not the best method, even if rates are constant because…

Calibrated Strict Clock

$r\,t_2$

$r\,t_1$

? mya   1 mya   0

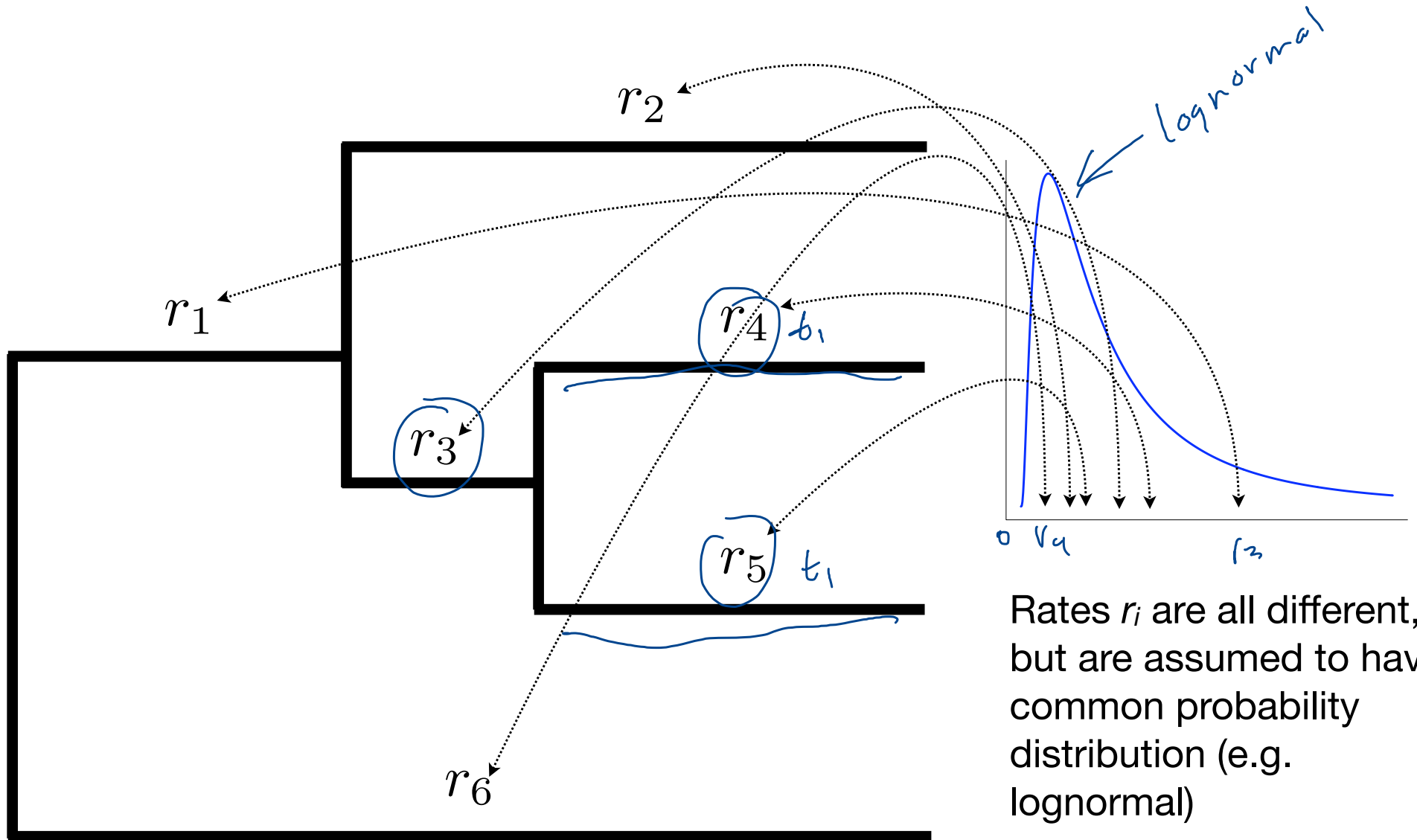…fossils seldom represent a direct ancestor, so there is a branch of unknown length throwing a wrench into the works.

Strict clock questionable

$$r_2 t_2$$

$$r_1 t_1$$

? mya   1 mya   0

But perhaps even worse is that for most data sets a strict molecular clock can be ruled out, so $r_1 \neq r_2$.
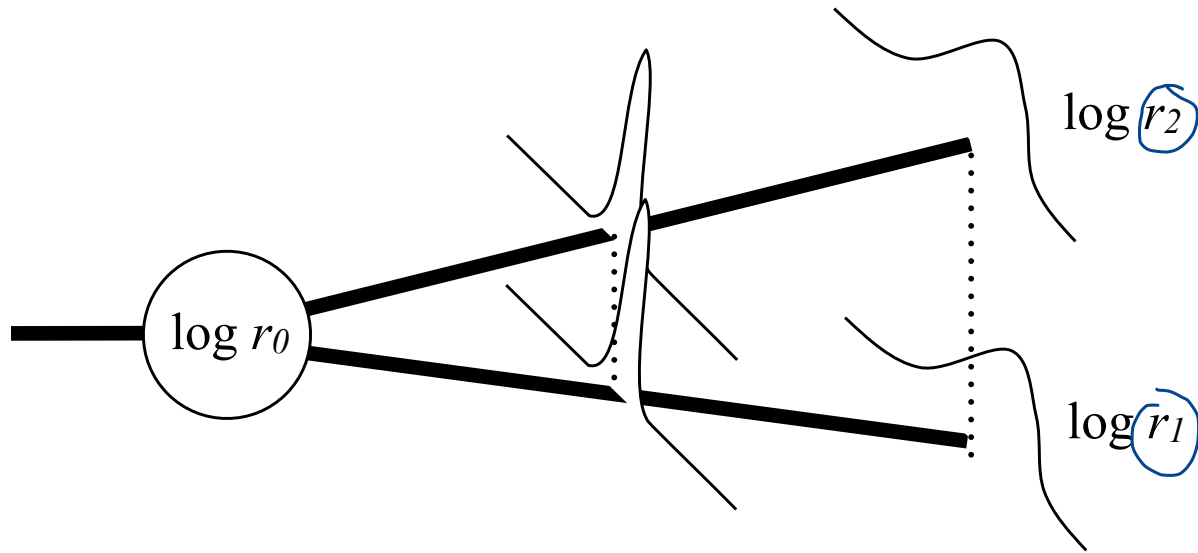
Uncorrelated relaxed clock model

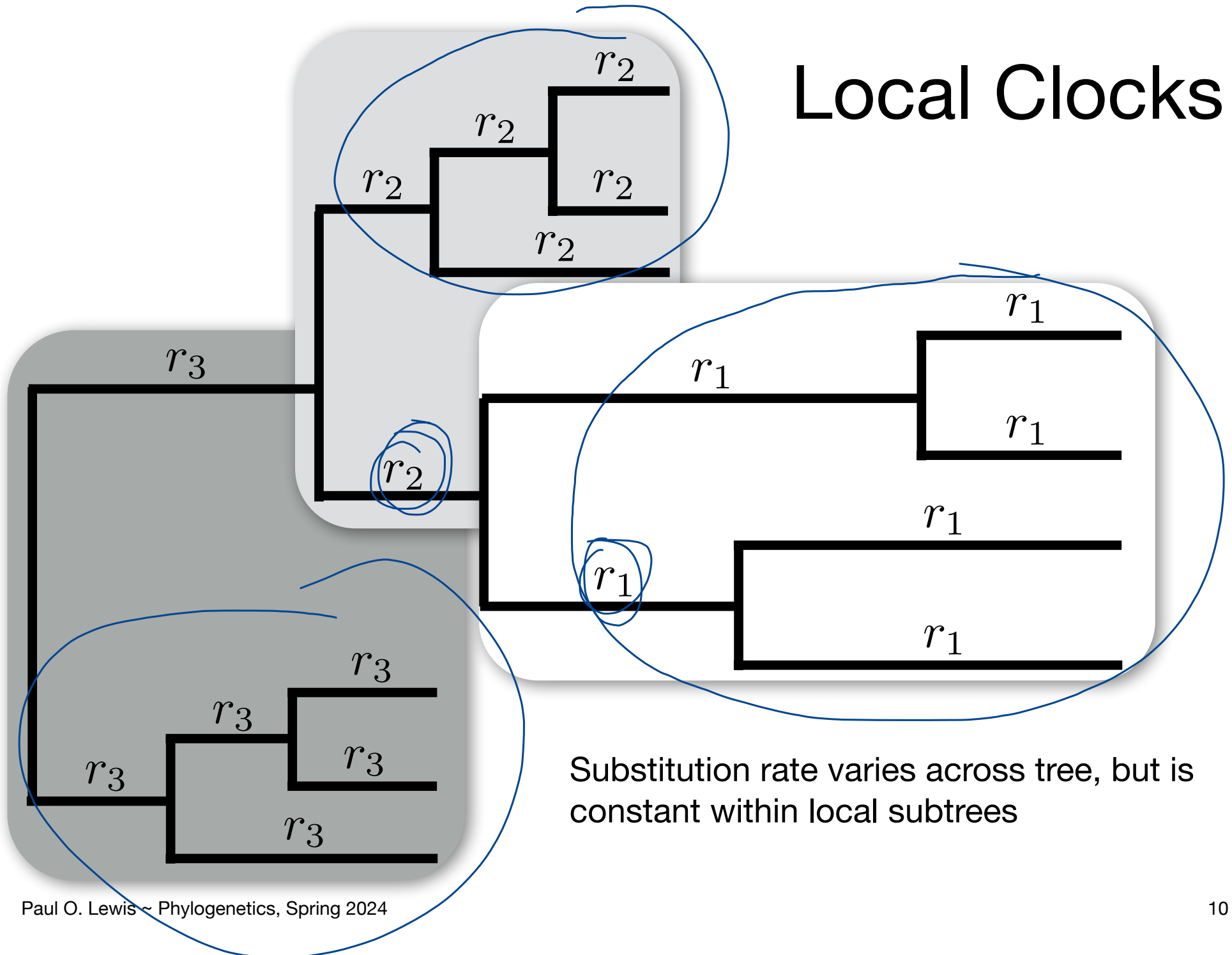Rates r_i are all different, but are assumed to have a common probability distribution (e.g. lognormal)

Correlated relaxed clock model

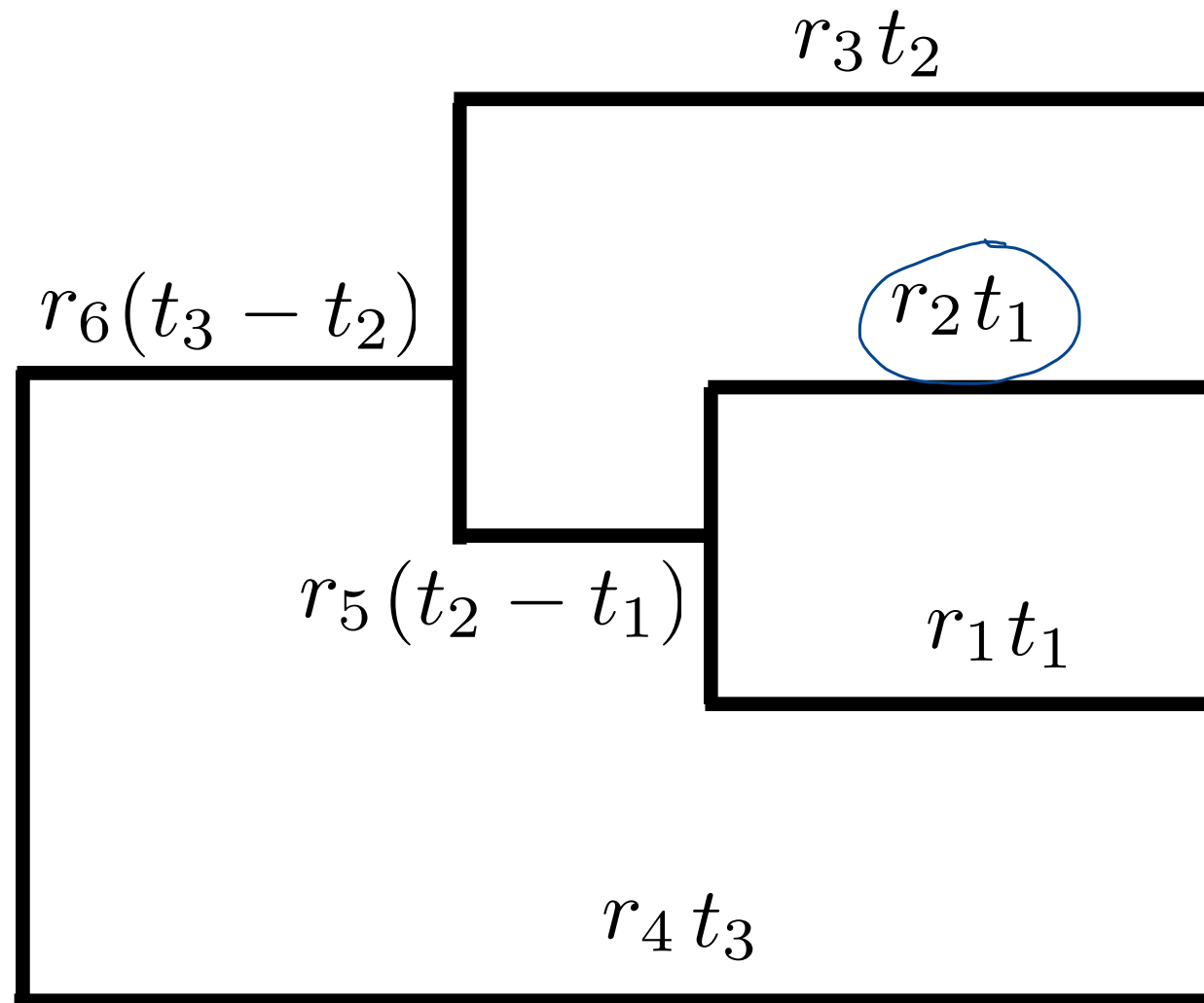…or are assumed to have diverged from a common ancestral rate.

Local Clocks

Substitution rate varies across tree, but is constant within local subtrees

Bayesian uncorrelated lognormal model

Drummond et al. 2006

Must know rates and times to compute likelihood

$r_3 t_2$

$r_6 (t_3 - t_2)$

$r_2 t_1$

$r_5 (t_2 - t_1)$

$r_1 t_1$

$r_4 t_3$

The goal: credible intervals for divergence times

Assume fixed (rooted) tree topology for now

The goal: credible intervals for divergence times

STOPPED HERE

2024 - 04 - 11

height