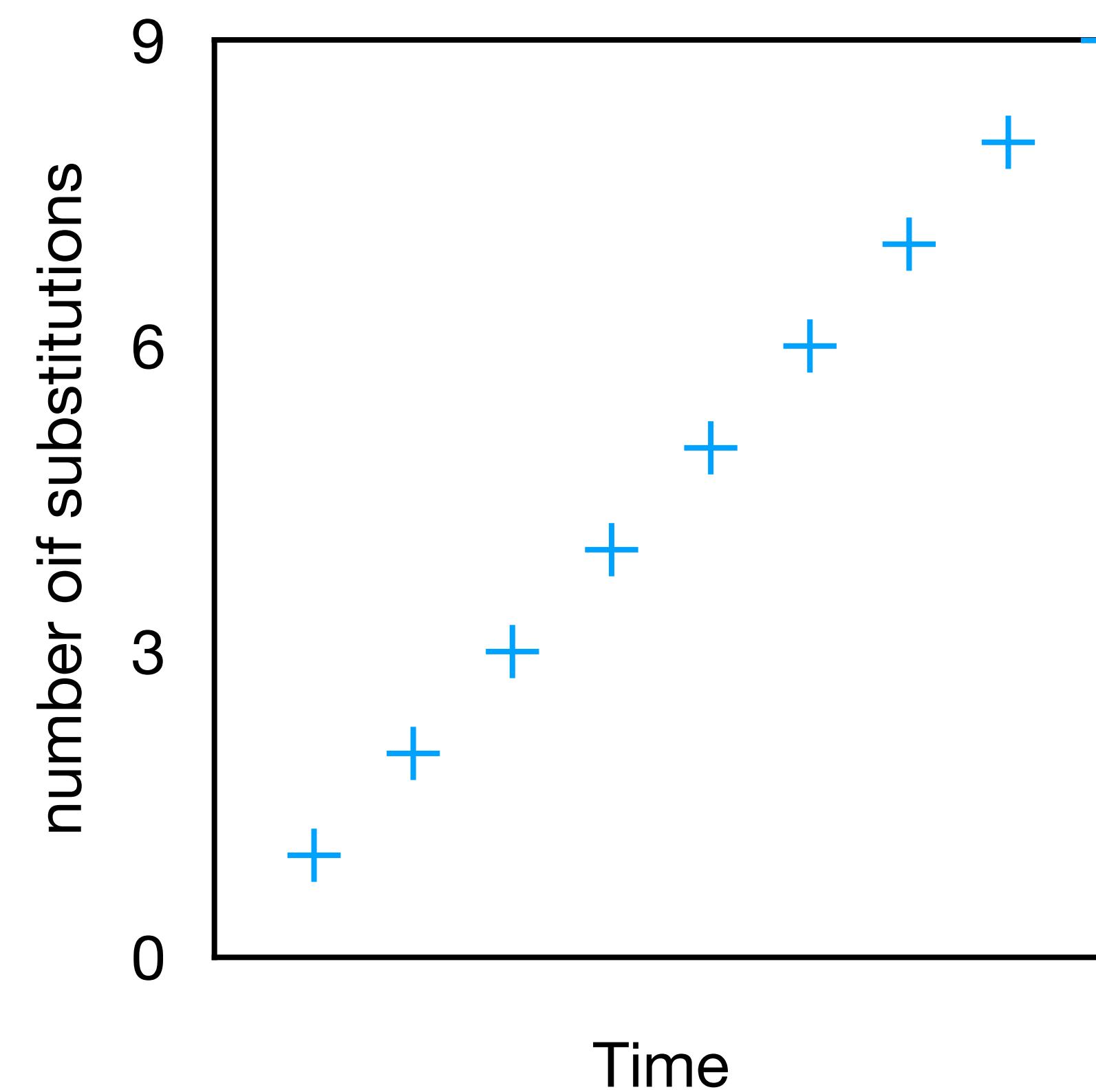
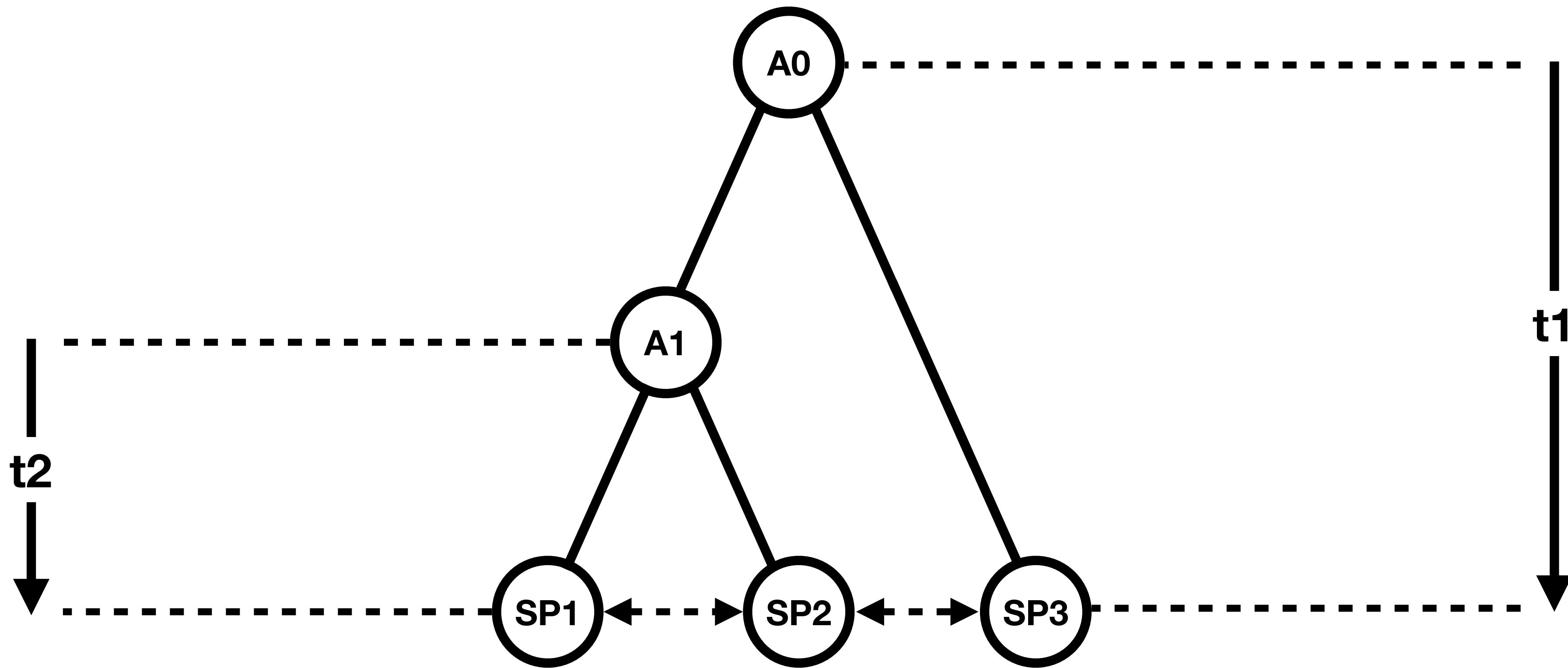


divergence time
analyses

The molecular clock hypothesis suggests that **genetic changes** tend to accumulate at a relatively **constant rate** over long periods 

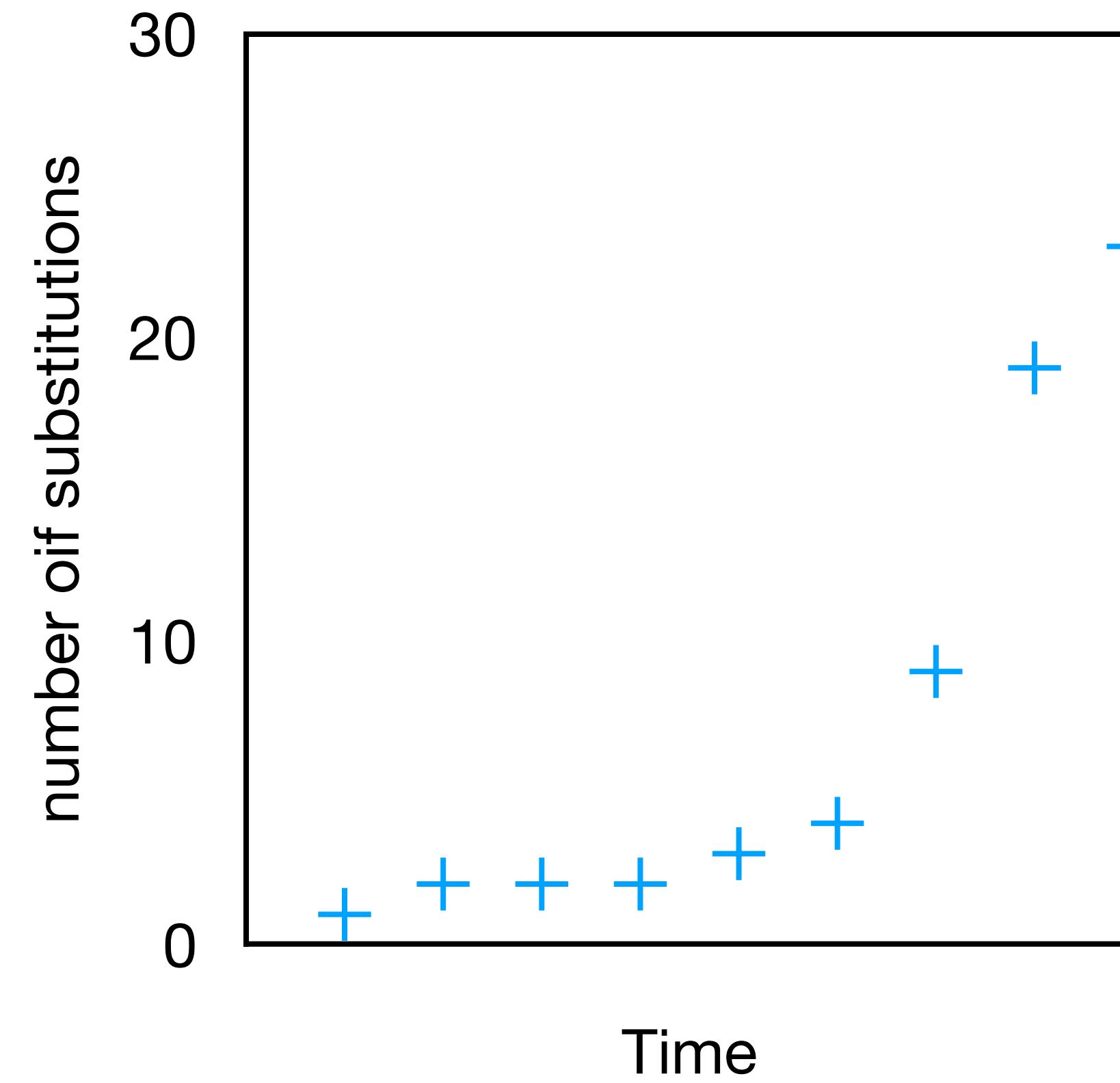


That would make estimating divergence times between species a breeze 😊



$$\text{Divergence Time } (\text{sp}_1 - \text{sp}_3) = \frac{\text{Molecular Distance } (\text{sp}_1 - \text{sp}_3)}{2}$$

... however, **the molecular clock assumption is often violated** ...



... here, the plot suggest a non-linear pattern, suggesting accelerated rates at some point of the evolutionary history of a lineage ...

 **Strict Molecular Clock**

assumes **constant rate of evolution** across all lineages
often **unrealistic** for large or heterogeneous datasets
can produce **biased age estimates** when rate variation exists

From the **relaxation** of the strict molecular clock **assumption** we get ...

 **Local Molecular Clock**

different **substitution rates** apply to different **monophyletic groups** within a tree
each **local clock group** evolves at its own constant rate

 **Relaxed Molecular Clock**

allows **rate variation among lineages**
2 main types:

- **Uncorrelated clocks:** rates vary independently (e.g., lognormal or exponential priors)
- **Correlated clocks:** rates in neighboring branches are related



Autocorrelated Relaxed Clock

- Substitution rates change gradually along the tree
- Daughter branches tend to inherit their parent's rate
- Models evolutionary “clock drift” (e.g., lognormal distribution)



Uncorrelated (General) Relaxed Clock

- Substitution rates are drawn independently for each branch
- No assumption of rate similarity between parent and daughter
- Common distributions: lognormal, exponential, gamma

Least Squares Dating (LSD)

- fits divergence times to branch lengths by **minimizing squared differences**
- uses a **strict or relaxed clock**, depending on rate smoothing
- requires a fixed topology and branch lengths, extremely fast and ideal for large datasets
- less sensitive to model violations, but lacks full probabilistic framework
- software: IQ-TREE

Penalized Maximum Likelihood (PML)

- **semi-parametric** approach balancing fit and rate smoothness
- allows rate heterogeneity while discouraging sudden shifts
- requires a **fixed topology and branch lengths**
- computationally efficient for large-scale phylogenies
- software: r8s, treePL

Bayesian Inference

- fully **probabilistic framework**: models uncertainty in topology, rates, and dates
- can implement the **Fossilized Birth-Death (FBD) model** including speciation, extinction, etc
- avoids arbitrary calibration bounds by integrating fossil distributions
- can implement **total-evidence dating** combining morphological + molecular data
- software: BEAST, RevBayes

Tip Dating

- dates are assigned directly to the tips of the tree
- uses dated samples (e.g. ancient DNA or fossils as terminals)
- integrates “extinct” information as part of the phylogenetic inference
- allows estimation of divergence times without node calibrations
- more sensitive to calibrations
- often used in total-evidence dating approaches

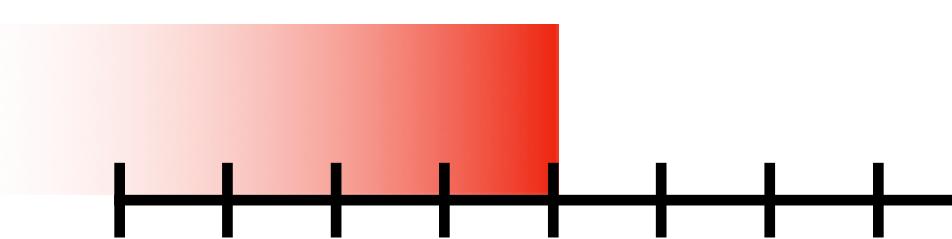
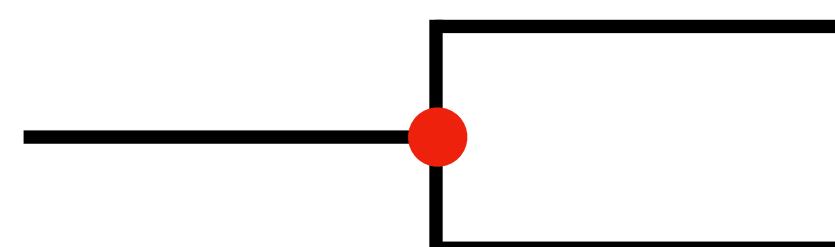
Node Dating

- dates are assigned to internal nodes of a phylogeny
- Relies on fossil calibrations placed at ancestral nodes
- Fossils provide minimum age constraints
- Assumes correct placement of fossils and accurate tree topology
- Common in analyses using only extant taxa

fossils

...

usually give only the minimum age constraint



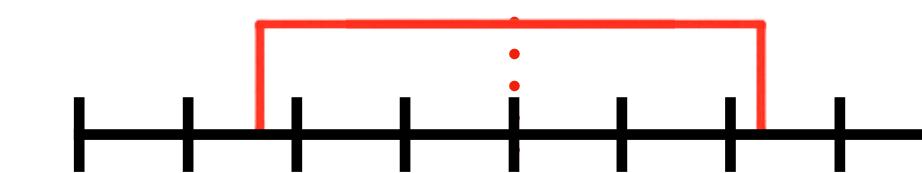
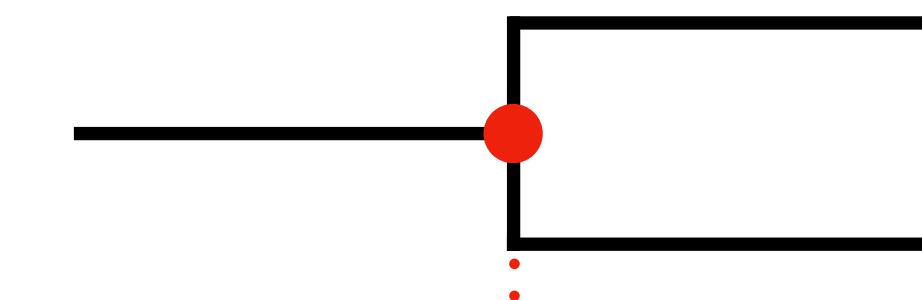
MAXIMUM
(SOFT)

MINIMUM
(HARD)

more common fossil prior shapes

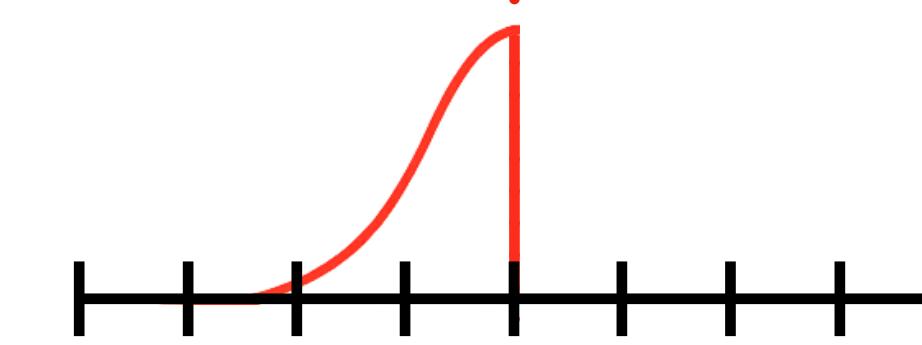
Uniform

- equal probability between minimum and maximum ages
- dimple and non-informative, often unrealistic
- use only when maximum bound is well-justified



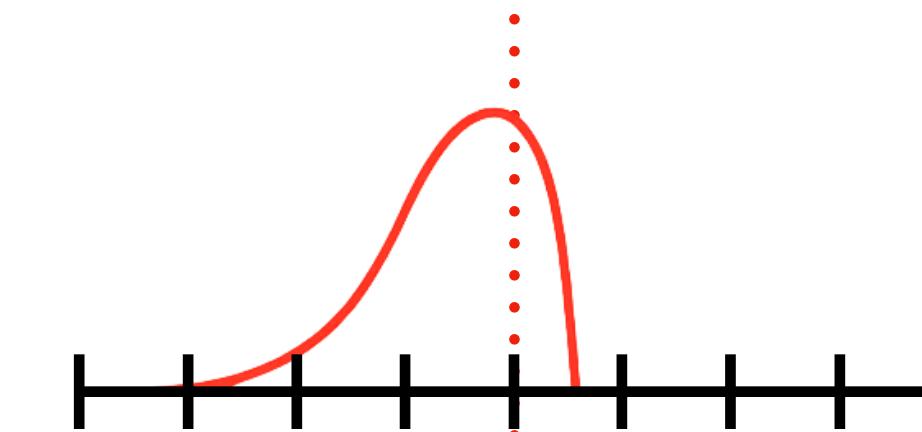
Lognormal

- probability concentrated just above the minimum
- heavy tail toward older ages
- divergence likely predicated the fossil, but not arbitrarily far back



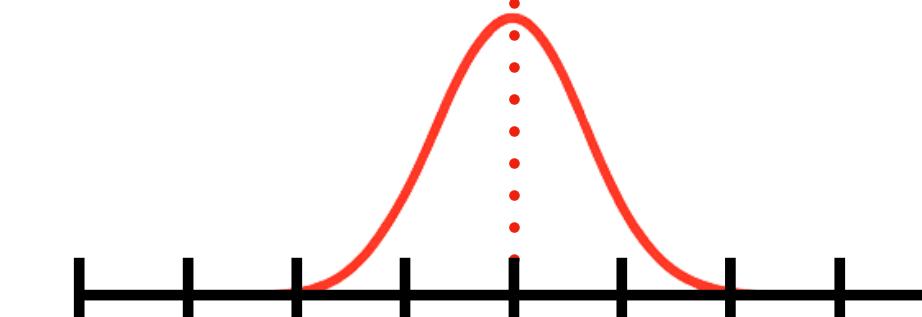
Exponential

- peaks at the minimum bound, then declines rapidly
- suitable when fossil is likely close to the actual divergence
- simple and interpretable with one parameter (mean)



Normal

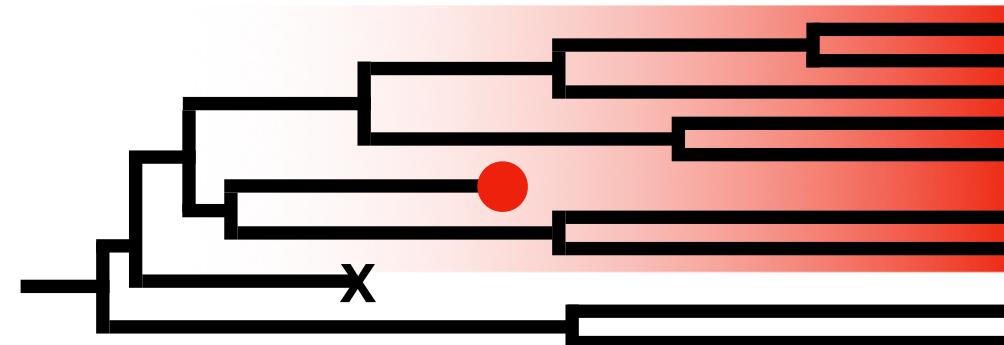
- bell-shaped, centered at a hypothesized mean
- requires strong justification for both center and spread
- less common for fossils, more for geological calibrations



fossils placement: Crown versus Stem

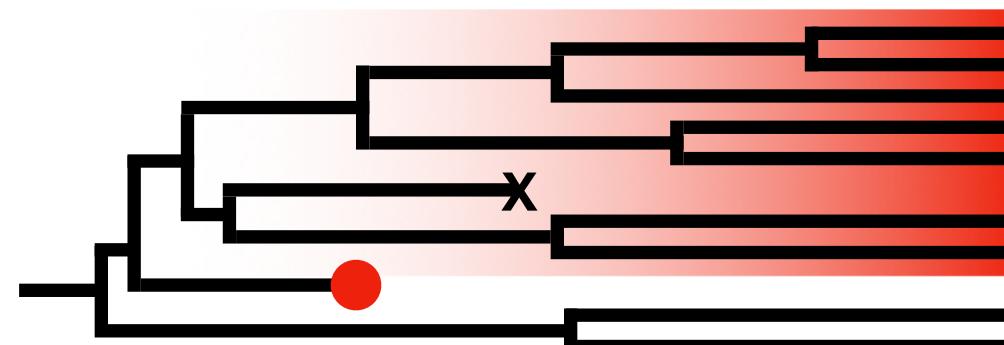
Crown Group

- includes the last common ancestor of all living members of a clade and all of its descendants.
- fossils placed here are within the clade, sharing all its defining synapomorphies.
- Exp: a fossil assigned to crown birds must postdate the divergence from non-avian dinosaurs.



Stem Group

- includes extinct lineages closer to a crown group than to any other, but outside the crown itself.
- stem fossils help bracket the minimum age of the crown group.
- Example: feathered theropod fossils - e.g. Archaeopteryx - are stem birds.



FINISH