Molecular clock: insights and pitfalls

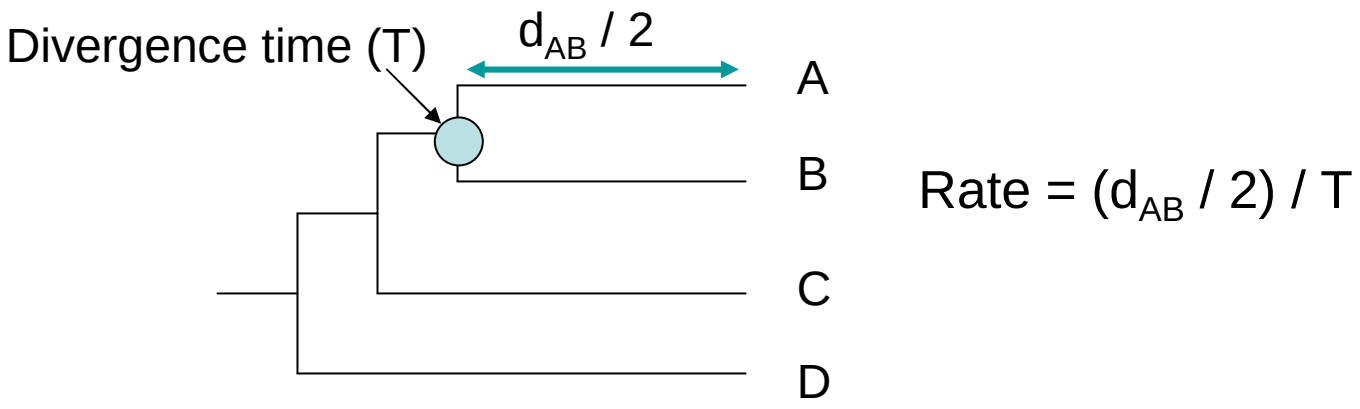
- An historical presentation of the molecular clock.
  - Birth of the molecular clock
  - Neutral Theory of evolution
  - The Nearly Neutral Theory of Evolution
- The molecular clock is a « sloppy » clock.
  - A variable « tick rate »
  - Different rates for different groups
- Calibrating a molecular phylogenetic tree
- Dating Methods

A historical presentation of the molecular clock

-1951-1955: Sanger sequenced the first protein, the insulin.

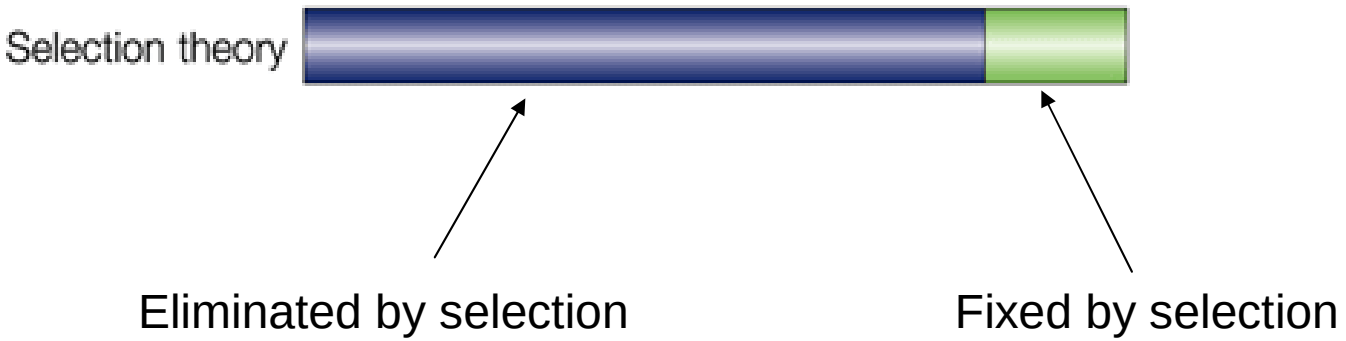
Give the potential for using molecular sequences to construct phylogeny

-1962: Zuckerkandl and Pauling: AA differences between the hemoglobin of different species is correlated with the time passed since they diverged.



Rate = (d_{AB} / 2) / T

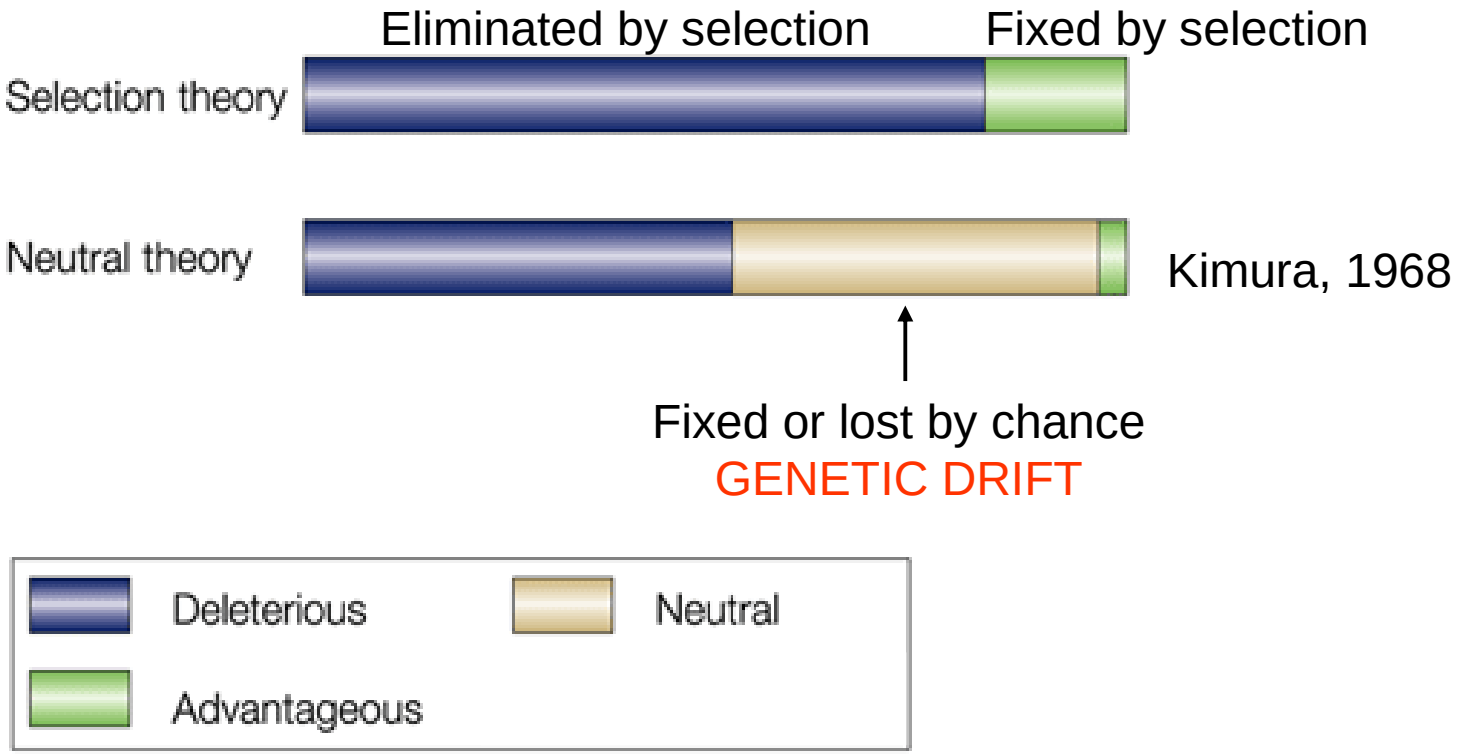
A historical presentation of the molecular clock



Selection theory more in agreement with the rate of morphological evolution

(Modified from Bromham and Penny, 2003)

The Neutral Theory of molecular evolution



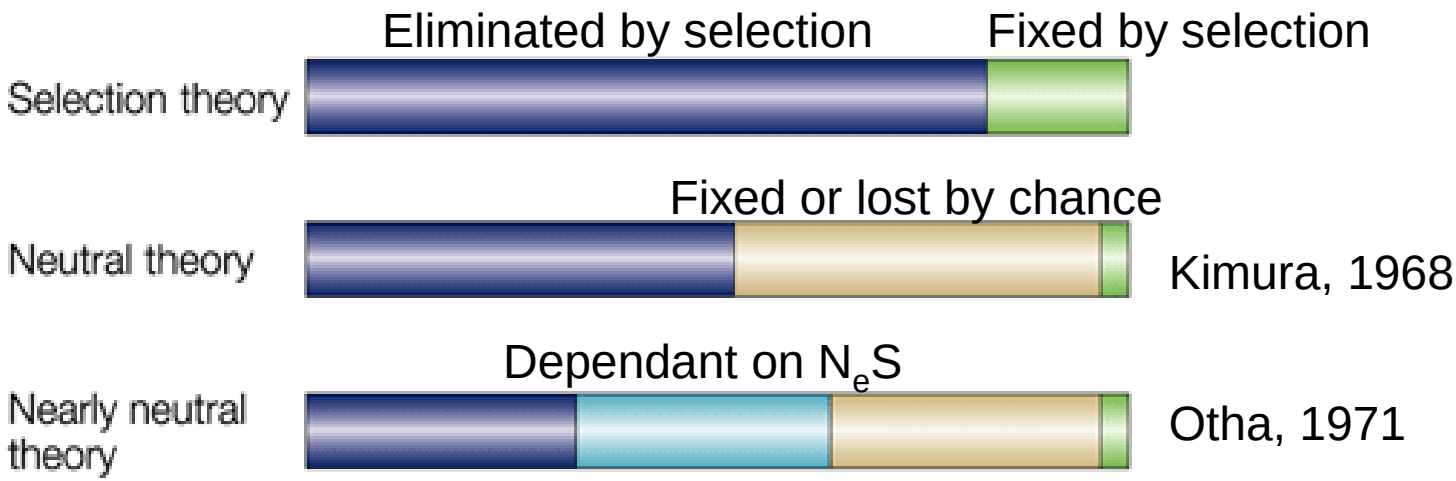
$2N_e\nu * 1/2N_e = \nu$

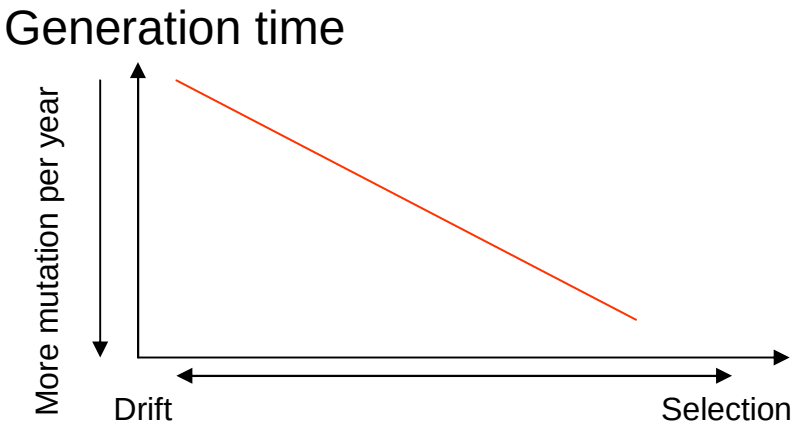
Molecular rate of evolution = Mutation rate

(Modified from Bromham and Penny, 2003)

The Nearly Neutral Theory of molecular evolution

Chronological time versus generation number





$\nu / (gN_e) \sim \kappa$

(Modified from Bromham and Penny, 2003)

A historical presentation of the molecular clock

In the early 1970's, DNA started to be investigated by the mean of DNA-DNA hybridization techniques:

•Laird et al. (1969): mutation rate (per year) 10 times higher between murids and artiodactyles.

•Kohne (1971): lower rate of DNA mutation in human than in great apes.

In the 1980's, development of DNA sequencing techniques:

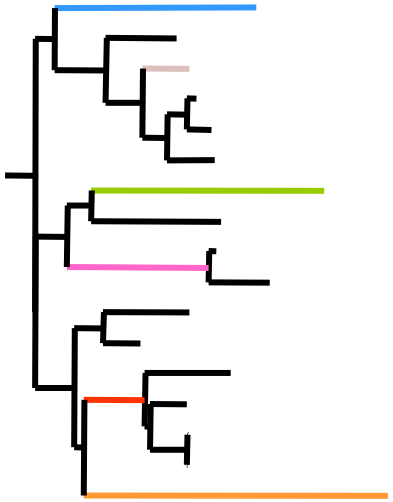
• Wu and Li (1985): large mutation rate difference between murids and human.

Evolutionary rate does not look constant anymore

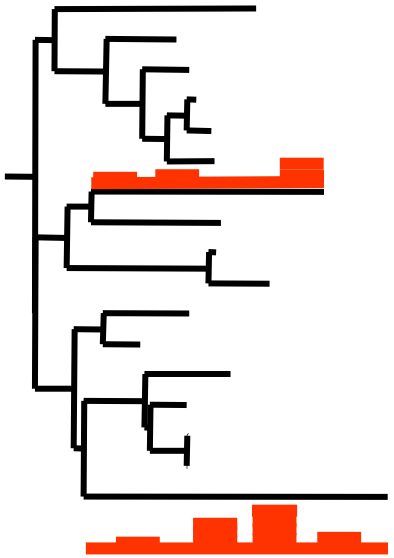
The molecular clock is a « sloppy » clock.

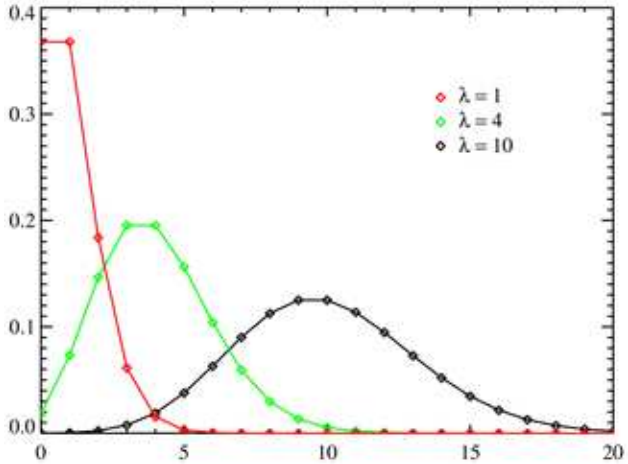
$$2N_e\nu * 1/2N_e = \nu$$

‘Lineage effect’
Variation on substitution rate between lineages



‘Residual effect’
Unevenness of substitution rate in a lineage.
Molecular clock probabilistic not deterministic.





Overdispersed poisson distribution

The molecular clock is a « sloppy » clock.

Changing the balance between selection and drift:

- Covarion model of protein evolution.

  Selective pressures on an amino acid or nucleotide site change through time because codons mutation rate varies in an autocorrelated manner due changes in functional constraints.

- Site in a gene can undergo a change in selection pressure.

  Burst of AA changes as a protein is evolving into a new role (cf. herbivory)

- Entire gene can undergo a change in selection pressure.

  Pseudo-genes or duplicated copy of a gene evolving into a new role.

- Effect of population size.

  Reduction in population size: burst of fixation of nearly neutral alleles.

'Lineage effect':

- Difference of repair equipment between taxa (biochemical factor).

  Mutation during replication and damaged not repaired.

- Species with higher metabolic rate might incur more DNA damage.

  High metabolic rates: higher concentration of mutagens and therefore higher mutation rate.

- Generation time.

  Mutation rate linked with germ-cell division in each generation.

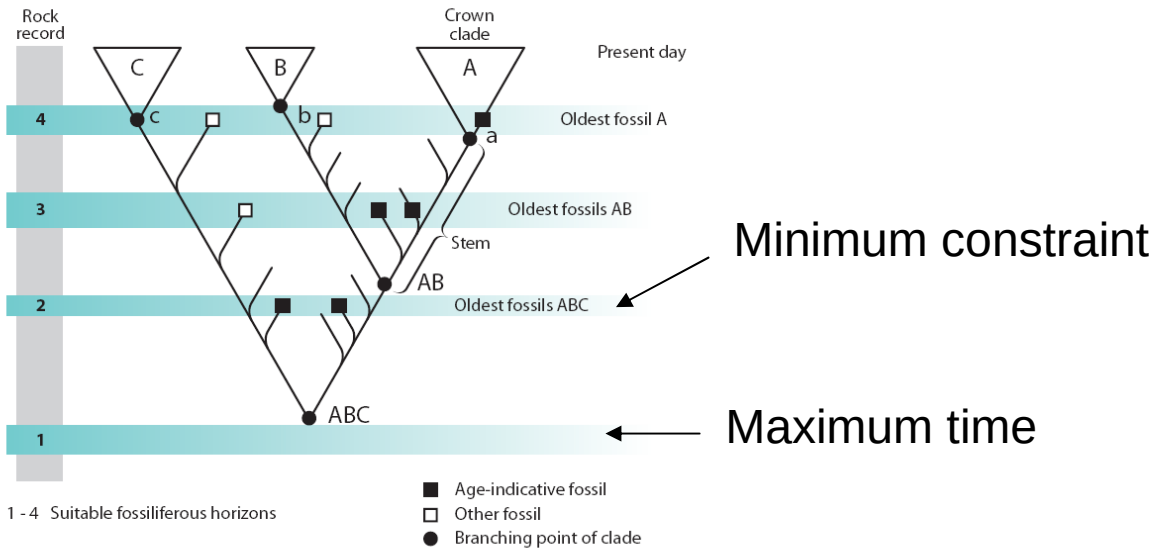
Calibrating the molecular tree

Uncertainty linked to the paleontologic record:

- Identifying the correct position of a fossil in a tree

- The fossil record is incomplete: the most ancient date at which two lineages can be detected represents only the minimum time back to their common ancestor.

- There is an error associated with the process of dating fossils (radioisotope dating)



Calibrating the molecular tree

How can the uncertainty be taken into account ?

- Use as much fossil calibration as possible.

- Calibration points (fossils ages) are not ‘hard bounds’ but ‘soft bounds’ / upper and lower bound can be used.

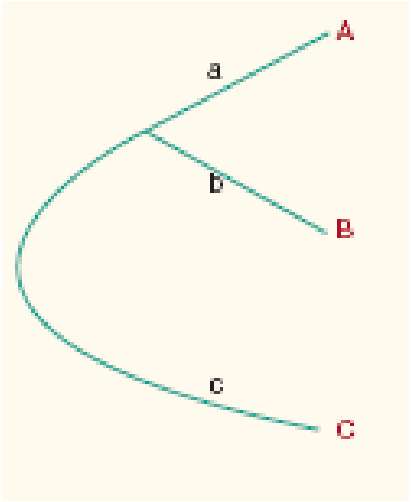
- Testing the reciprocal cross-validity of each calibration point (Near and Sanderson, 2004).

Methods and programs

The relative rate test: testing the global molecular clock (Fitch, 1976)

Known distances a+b, a+c, b+c

From these distances a and b can be deduced and then compare with a χ^2 .



Outgroup

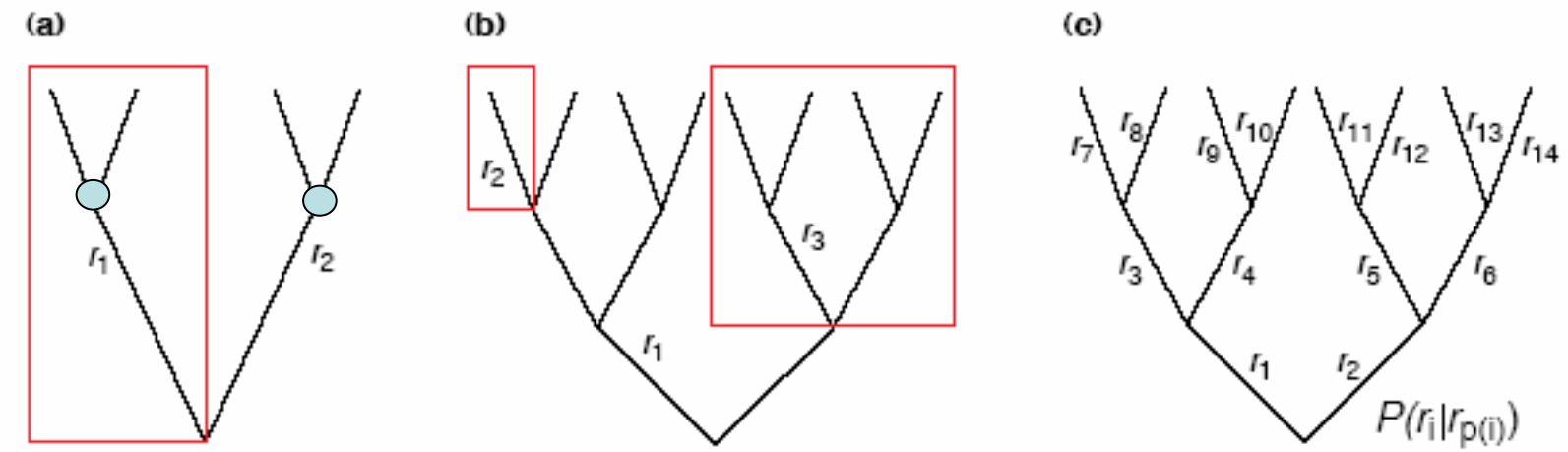
↓

Remove the species with an significantly different rate of evolution

↓

Apply the global molecular clock

Methods and programs



Quartet Dating

Local clocks

Relaxed clocks
(rate smoothing)
Autocorrelation of rates

↓

Contradicted by Drumond et al. (2006)

Age estimates: how wrong can they be ?

Probably not quite true but:

-Increase in sequence length + marker number (genomic scale)

- Increase of taxon sampling

- Increase of the accuracy of fossil age estimates

- Improvement of the methods modeling molecular evolution

Its getting better…

The Neutral Theory of molecular evolution

Rate of evolution: rate at which new alleles (created by mutation) are substituted for other alleles already present in the population.

Kimura, 1968: When genetic drift is the only force in action

Molecular rate of evolution = Mutation rate

ν: Rate of neutral mutation per allele per unit time

N_e : Effective population size

$\nu 2N_e$: Rate of creation of new neutral alleles per generation

$1/2N_e$: Chance of a new allele to be fixed

$$2N_e\nu * 1/2N_e = \nu$$

The different rate of evolution between proteins could be explain by the proportion of sites at which mutations are neutral.