E-content

M.Sc. Zoology (Semester II)
CC8- Biochemistry

Unit: 3.3b

Molecular Clock

Dr Gajendra Kumar Azad
Assistant Professor
Post Graduate Department of Zoology
Patna University, Patna
Email: gkazadpatnauniversity@gmail.com

The tempo and mode of evolution are central themes of biological research. This places importance on the estimation of evolutionary timescales, which provide the backdrop for our interpretations of evolutionary patterns and processes.

Traditionally, such inferences were made from the fossil record, coupled with radiometric dating. Fossils can provide an estimate of when different lineages first appeared and when species diverged from each other. In many cases, however, such data are unavailable, forcing us to look elsewhere for a source of temporal information.

The “molecular clock,” proposed in the 1960s (***Zuckerkandl and Pauling***, 1962, 1965), allows evolutionary timescales to be estimated using genetic data

Molecular clock is abased on a hypothesis that predicts a constant rate of molecular evolution among species. It is also a method of genetic analysis that can be used to estimate evolutionary rates and timescales using data from DNA or proteins.

The molecular clock arises from the observation that the amount of difference between the DNA of two species is a function of the time since their evolutionary separation. This provides a universal tool not only for placing past evolutionary events in time, but also for exploring the mechanisms and processes of evolution.

Discovery of the molecular clock

An early and unexpected finding from the molecular revolution was the discovery that a given protein has a characteristic rate of evolution, but that genes differ in their characteristic rates.

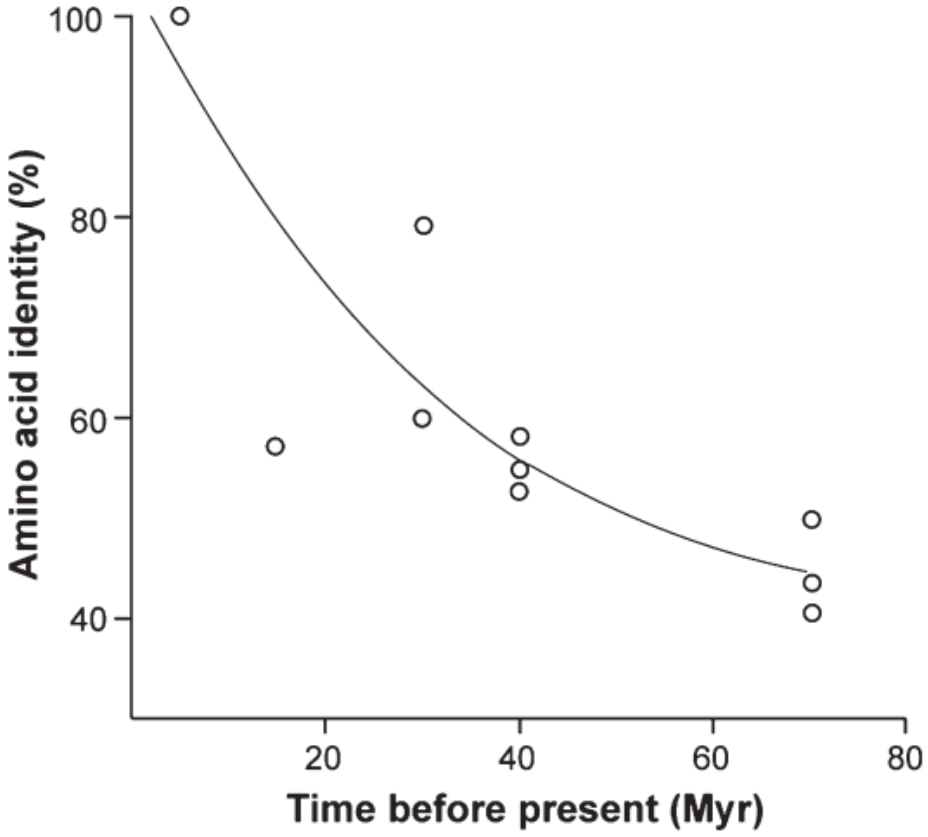
The molecular-clock hypothesis was put forward by Emile Zuckerkandl and Linus Pauling (1962), who assumed a constant evolutionary rate in their analysis of globin proteins from vertebrates.

They observed about 18 differences in amino acids between horse and human and estimated the mutation rate by assuming that the divergence between these two species occurred 100–160 Ma (mega annum or millions of years) ago.

Upon extrapolating this rate, Zuckerkandl and Pauling (1962) estimated that humans diverged from gorillas about 11 Ma ago. They also estimated that different copies of globin genes first diverged from each other in the late Precambrian. They reported a strikingly linear rate of accumulation of amino-acid differences over evolutionary time.

In the following years, further studies produced evidence of clocklike evolution in other proteins. Doolittle and Blomback (1964) found a simple relationship between sequence identity and time since divergence in mammalian fibrinopeptides (Figure).

Figure: Plot of amino acid identity of fibrinopeptides from various pairs of mammals, plotted against time since divergence. The divergence times are based on estimates from the fossil record (Data are from Doolittle and Blomback (1964)).



A year later, Zuckerkandl and Pauling (1965) coined the term “molecular evolutionary clock.”

Other proteins also showed a constant rate of molecular evolution across species, but with each protein having a different characteristic rate: histones were exceptionally slow, cytochrome *c* was faster (but slower than haemoglobins) and fibrinopeptides were faster still.

This relative equality of evolutionary rates for a given protein was unexpected; it had been assumed that, as with morphological evolution, there would be large variation in the rate of change both between species and over evolutionary time.

Criticisms of the molecular clock were partly motivated by the apparent lack of uniformity in the pace of morphological evolution, which was presumably linked to molecular evolution.

Meanwhile, there was growing evidence of high evolutionary rates in various proteins, suggesting that a large proportion of the changes in amino acids must have a negligible impact on evolutionary fitness.

As a response, Motoo Kimura (1968) proposed the neutral theory of molecular evolution, which states that many mutations have such a small effect on the fitness of an organism that they can be considered as “neutral.”

This can be explained by the fact that many amino acids in a protein can be exchanged for other amino acids with similar biochemical properties, with negligible impact on the overall function or structure of the protein.

This differed from previous models that had considered that most amino-acid changes would be either favourable (positively selected) or deleterious (removed by selection).

Kimura reasoned that advantageous mutations would be relatively rare, deleterious mutations would be rapidly removed by selection and that a large proportion of possible amino acid changes would have no practical effect on the functioning of the protein.

The rate of accumulation of these neutral mutations would be influenced only by the mutation rate, and so would be relatively constant, as long as the base mutation rate remained unchanged.

This is a fundamental result; it predicts that the long-term rate of neutral molecular evolution in species is the same as the neutral mutation rate in individuals.

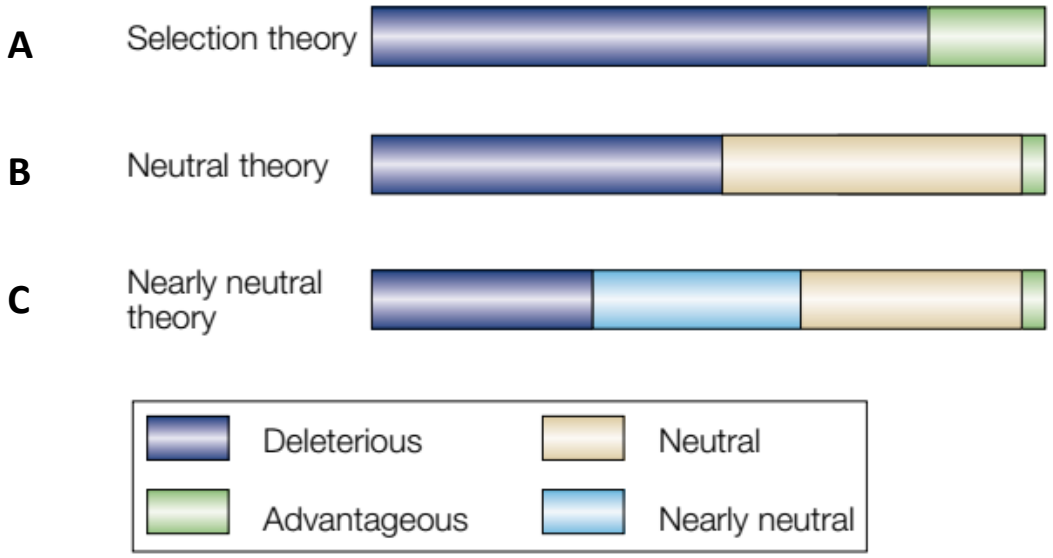
One of the predictions of the neutral theory is that rates of molecular evolution are constant among lineages. However, this prediction refers specifically to the rate of genetic change per generation. As a consequence, we expect to see a generation-time effect, whereby species with shorter generations tend to evolve more quickly per unit of time.

For example, a higher rate would be observed in rodents than in whales. This is based on the assumption that most mutations occur during the replication of germline DNA. This was in contrast with the rate of protein evolution, which appeared to be independent of generation time.

In response to the shortcomings of the neutral theory, Tomoko Ohta (1972, 1973) proposed the nearly neutral theory of molecular evolution. In this framework, there is a large class of "**nearly neutral**" mutations that have small effects on an organism's fitness.

In contrast with the results of the neutral theory, the nearly neutral theory states that the population sizes of species have a significant influence on the molecular evolutionary process.

Selectionist, neutral and nearly neutral theories



A: Selectionist theory: early neo-Darwinian theories assumed that all mutations would affect fitness and, therefore, would be advantageous or deleterious, but not neutral.

B: Neutral theory: the neutral theory considered that, for most proteins, neutral mutations exceeded those that were advantageous, but that differences in the relative proportions of neutral sites would influence the rate of molecular evolution (that is, more neutral sites would produce a faster overall rate of change).

C: Nearly neutral theory: the fate of mutations with only slight positive or negative effect on fitness will depend on how population size affects the outcome.

The molecular clock is a ‘sloppy’ clock

The neutral theory predicts that, for a given mutation rate and proportion of neutral sites, the rate of molecular evolution should be constant.

This constancy of rates is broadly supported by observation: the amount of divergence between genes tends to increase with the time since evolutionary separation, and in many cases the increase seems linear, as in the classic example of globin genes.

In a more recent example, the nucleotide distance between sister species on Hawaiian islands, plotted against geological estimates of island age, gave impressively linear relationships for both birds and fruitflies. However, empirical studies have also shown a great deal of variation in the rate of molecular evolution.

The neutral theory allows for two sources of rate variation in the molecular clock: the ‘sloppiness’ of the ‘tick rate’ and changes in the mutation rate. These types of rate variation do not necessarily arise from different mechanisms.

However, they do give rise to two types of error in molecular date estimates that contribute to ‘residual effects’ (unevenness of substitution rate in a lineage) and ‘lineage effects’ (variation in substitution rate between lineages) on the rate of molecular evolution.

A variable ‘tick rate’

The molecular clock is probabilistic, not deterministic — it ticks at irregular intervals. This distribution of intervals between substitution events (‘ticks’) — commonly described by a POISSON DISTRIBUTION — adds large confidence intervals to date estimates because the time taken to produce the observed substitutions cannot be known precisely.

In fact, the molecular clock is generally even sloppier: observations indicate that the pattern of substitution intervals for many genes has a broader distribution, which might be described as an ‘overdispersed’ Poisson distribution.

So, where does this extra variation come from?

Many processes could affect the number of substitutions per unit time, primarily by changing the balance between the relative influences of selection and drift, either for specific genes or sites in genes, or across whole genomes.

Another source of variability comes from the effect of population size on the rate of fixation of mutations.

Tomoko Ohta extended the neutral theory by recognizing the critical role of effective population size.

Small populations are more severely affected by stochastic fluctuations in allele frequencies, so genetic drift can overpower selection for alleles with small selection coefficients.

So, if a population undergoes a marked reduction in population size due to an environmental catastrophe, this event might be accompanied by a burst of fixation of nearly-neutral alleles.

In this way, population fluctuations might add to the sloppiness of the molecular clock.

How reliable is the molecular clock?

The molecular clock allows a valuable insight into the biochemical, selective and population processes that underlie genetic evolution.

It also provides a remarkable tool for investigating evolutionary history: if the rate of molecular evolution is relatively constant, then the amount of genetic difference between two species gives a measure of the time since their evolutionary separation .

This molecular timescale can provide insights into the history of all organisms from which we can obtain genetic sequences.

This is valuable in the case of organisms with no fossil record — such as viruses — or for which the sampling of the fossil record is patchy in time or space.

Because of the universality of DNA (or RNA), molecular clocks can reach all timescales of evolution, from population divergences to the last common ancestor of the five kingdoms of life.

But this apparently simple technique has resulted in some controversial molecular dates. Some seem to contradict other lines of evidence, such as the study that produced a molecular date estimate for the split between kingdoms that is markedly younger than the earliest fossils.

Other molecular dates have been used to challenge the reliability of the fossil record, and engendered debate over the tempo and mode of macroevolution, such as the discrepancy between molecular and fossil dates for the origin of animal phyla.

Molecular evolutionary theory leads us to expect two types of error in molecular clock estimates.

First, the sloppy nature of the substitution rate results in large variance around the amount of genetic difference expected for any given time period, adding a large degree of imprecision to molecular date estimates.

Second, the nearly neutral theory predicts that the rate of molecular evolution is influenced by mutation rate, population size and the relative proportions of sites with different selective coefficients; these factors might differ between genes, between species and over time, potentially resulting in consistent bias in date estimates.

Controversies of molecular clock

The molecular clock has had a long history of controversy, beginning with the criticisms of the assumption of rate constancy and continuing with the debate over the relative merits of the neutral theory compared with natural selection.

Current debates include the use of calibrations in genetic dating analyses and the constancy of evolutionary rates across timescales.

The choice of calibrations has a substantial influence on the outcome of a molecular-clock analysis. Inadequate modelling of calibrations can lead to highly misleading estimates of evolutionary timescales. Therefore, the use of multiple calibrations are recommended for generating the phylogenetic tree.

Ignoring the calibrations can lead to artificially precise estimates of evolutionary timescales. Instead, the uncertainty should be incorporated into the dating analysis so that it can be included in the resulting estimates. This can readily be done in a Bayesian framework, in which the user can choose prior distributions that reflect the degree of uncertainty in the parameters and node times.

There is growing evidence that estimates of rates depend on the timescale of observation. If the time dependence of molecular rates is not taken into account, evolutionary timescales can be under- or overestimated by an order of magnitude.

References

Evolution by Douglas J. Futuyma

***Nature Education Knowledge* 4(4):1 (2013)**

An Introduction to Genetic Analysis. 7th edition. Griffiths

---------- End ----------