E-content

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

Unit: 3.3a

Rates of molecular evolution

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

The molecular evolutionary rate measures the frequency with which DNA or protein sequence mutations are fixed (i.e., shared by most individuals) in a population. On the other hand, the mutation rate refers to the amount of change in a DNA or protein sequence for a given unit of time. These two intrinsically related processes have been key to our basic understanding of molecular evolution.

Absolute rates of molecular evolution vary among regions of the genome because of differing constraints. Functionally important genes often evolve at an extremely slow pace, because many new mutations are detrimental to the organism and are rapidly removed by natural selection.

As a consequence, such genes are often conserved even among distantly related organisms, making them useful genetic markers for studying deep evolutionary relationships.

For example, histones, which are proteins that play an important role in binding and packaging DNA, have a very low substitution rate. In contrast, non-coding DNA has fewer functional constraints and can evolve at a higher rate than protein-coding DNA.

The mitochondrial genome evolves rapidly in animals and experiences high rate of nucleotide substitutions (an order of magnitude higher than the average rate in the nuclear genome). Rates of change in the genomes of viruses are higher by several orders of magnitude. Furthermore, there is substantial variation in rates of molecular evolution across the tree of life.

The rate of molecular evolution can be predicted by using the molecular clock theory (Zuckerkandl & Pauling 1962). This theory predicts an increase in fixed amino acid mutations of protein sequences as a function of time. The molecular evolutionary rate is then expected to be constant, assuming that mutation rates are the same across the evolutionary lineages of a given protein.

The neutral theory of molecular evolution (Kimura 1983) explains this pattern by proposing that most mutations do not have an effect on the fitness of an organism.

Under this model, fixed DNA nucleotide mutations that do not result in changes to protein sequences are referred to as neutral or synonymous substitutions. Also, beneficial mutations are expected to be rare and deleterious ones quickly removed by natural selection.

What determines variation in the mutation rate and ultimately the molecular evolutionary rate? The answer to this question is debatable for a number of reasons. Evidence at the cell and molecular levels clearly supports the finding that mutations often derive from DNA replication errors or mutagens in the environment.

Organismal-level traits and population genetic processes, in turn, could influence the frequency of DNA replication errors and mutagen levels. Mutations are transmitted from generation to generation in the germ cell line of sexually reproducing organisms. Because DNA in germ cells replicates during meiosis (gamete differentiation), we expect short-lived species to have higher mutation rates.

For example, mice reproduce frequently and have short generation times. Their germ cell lines undergo more rounds of meiosis, thereby increasing the chances of DNA replication errors. Compared to long-lived species, mice also are expected to have larger populations with more individuals that are available to reproduce — that is, a larger effective population size.

Consequently, synonymous (i.e., neutral) mutations are more likely to occur, leading to higher molecular evolutionary rates. Furthermore, the larger effective population size promotes selection against non-synonymous mutations that reduce fitness, thereby removing them from the population.

Although it is possible that only one or a few mutations lead to a change in the specificity of a protein, the more usual situation is that DNA accumulates substitutions over long periods of evolution without making a qualitative change in the functional properties of the proteins that are encoded.

There may, however, be smaller effects influencing the kinetic properties, timing of production, or quantities of the encoded proteins that, in turn, will affect the fitness of the organism that carries them.

Mutations of DNA can have three effects on fitness.

First, they may be deleterious, reducing the probability of survival and reproduction of their carriers.

Second, they may actually increase fitness by providing increased efficiency or by expanding the range of environmental conditions in which the species can make a living or by enabling the organism to track changes in the environment.

Third, they may have no effect on fitness, leaving the probability of survival and reproduction unchanged; they are the so-called neutral mutations.

If a newly arisen mutation is effectively neutral then, there is a probability of $1/(2N)$ that it will replace the previous allele because of random genetic drift.

If the rate of appearance of new effectively neutral mutations at a locus per gene copy per generation is μ, then the absolute number of new mutational copies that will appear in a population of N diploid individuals is $2N\mu$.

Each one of these new copies has a probability of $1/(2N)$ of eventually taking over the population.

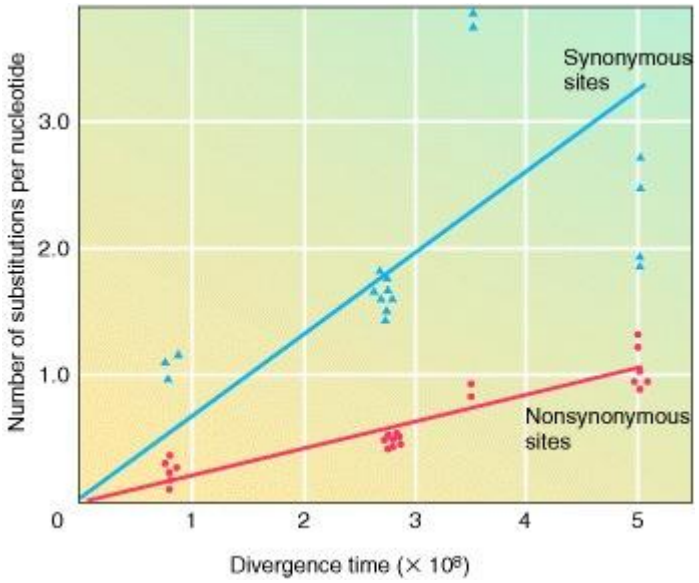
Thus, the absolute rate of replacement of old alleles by new ones at a locus per generation is their rate of appearance multiplied by the probability that any one of them will eventually take over by drift.

$$\text{Rate of neutral replacement} = 2N\mu \times 1/(2N) = \mu$$

That is, we expect that in every generation there will be μ substitutions of a new allele for an old one at each locus in the population, purely from genetic drift of effectively neutral mutations.

The constant rate of neutral substitution predicts that, if the number of nucleotide differences between two species is plotted against the time since their divergence from a common ancestor, the result should be a straight line with a slope equal to μ.

That is, evolution should proceed according to a molecular clock that is ticking at the rate μ. Figure shows such a plot for the β-globin gene. The results are quite consistent with the claim that nucleotide substitutions have been effectively neutral over the past 500 million years. Two sorts of nucleotide substitutions are plotted: synonymous substitutions that are from one alternative codon to another, making no change in the amino acid, and **nonsynonymous substitutions** that result in an amino acid change. Figure shows a much lower slope for nonsynonymous substitutions than that for synonymous changes, which means that the mutation rate to nonsynonymous substitutions is much lower than that to synonymous ones. This is precisely what we expect. The mutation rate to neutral alleles, μ, is the product of the intrinsic nucleotide mutation rate, *M*, and the proportion of all mutations that are neutral, *f*. That is,



$$\mu = M \times f$$

and it is reasonable that mutations that cause an amino acid substitution should more often have a deleterious effect, *s*, above the threshold for neutral evolution and therefore the proportion of neutral changes, *f*, will be smaller.

It is important to note that these observations do not show that synonymous substitutions have *no* selective constraints on them; rather they show that these constraints are, on the average, not as strong as for amino acids.

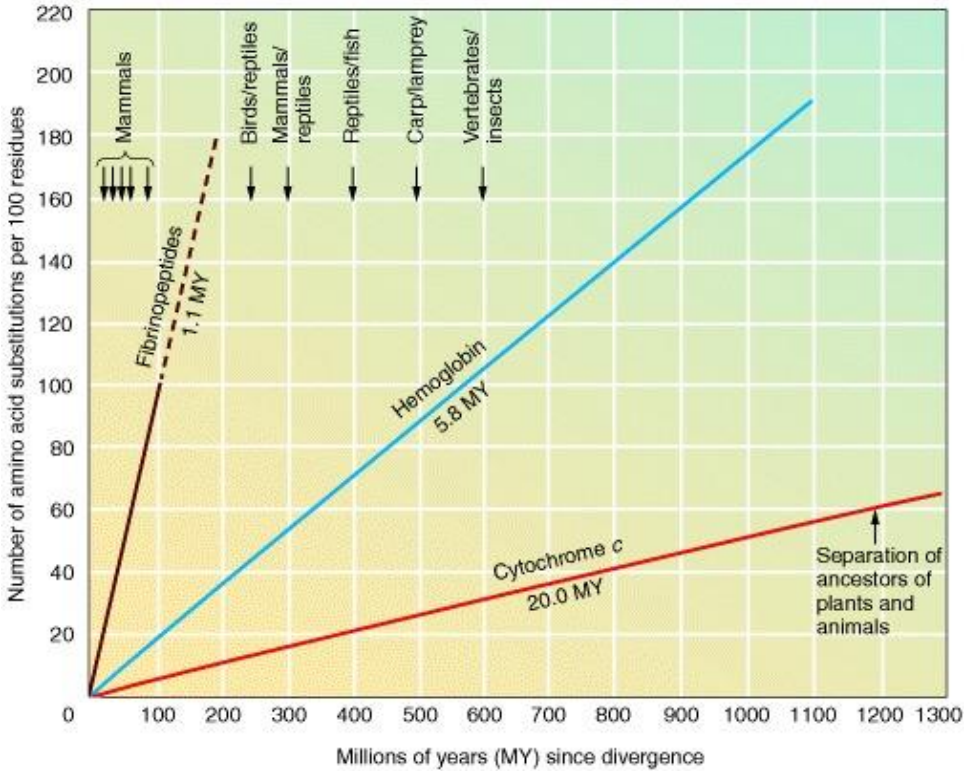
In fact, synonymous changes do have effects on probabilities of correct splicing, on the stability and lifetime of mRNA, on use by the translation apparatus of the available pool of tRNA molecules, and thus on the rate of translation and the folding of the translated polypeptide.

Another prediction of neutral evolution is that different proteins will have different clock rates, because the metabolic function of some proteins will be much more sensitive to changes in their amino acid sequence.

Proteins in which every amino acid makes a difference will have smaller values of the effectively neutral mutation rate, *Mf*, than will proteins that are more tolerant of substitution.

Figure shows a comparison of the clocks for fibrinopeptides, hemoglobin, and cytochrome *c*. That fibrinopeptides have a much higher proportion of neutral mutations is reasonable because these peptides are merely a nonmetabolic safety catch, cut out of fibrinogen to activate the blood-clotting reaction.

Figure: Number of amino acid substitutions in the evolution of the vertebrates as a function of time since divergence.



The three proteins—fibrinopeptides, hemoglobin, and cytochrome *c*—differ in rate because different proportions of their amino acid substitutions are selectively neutral.

The rate of neutral evolution for the amino acid sequence of a protein depends on the sensitivity of a protein's function to amino acid changes.

The demonstration of the molecular clock argues that most nucleotide substitutions are neutral, but it does not tell us how much of molecular evolution is adaptive.

One way of detecting adaptive evolution of a protein is by comparing the synonymous and nonsynonymous polymorphisms within species with the synonymous and nonsynonymous changes between species.

By comparing differences among a wide variety of organisms, biologists can test the prediction that DNA nucleotide sequences do indeed evolve according to a rate that, at least partially, depends on organism-level traits.

Generation time and metabolism, each to some degree or in combination, affect the mutation rates of some organisms and, thus, their molecular evolutionary rates.

Even so, some relationships among generation time, metabolism, and molecular evolution depend on whether the organism is a plant or an animal and the location of the genome within the cell (i.e., nuclear vs. organellar).

Also, differences in neutral vs. non-synonymous rates, when averaged together across long DNA sequences, could further complicate interpretations.

References

Evolution by Douglas J. Futuyma

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

An Introduction to Genetic Analysis. 7th edition. Griffiths

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