RATE OF MOLECULAR EVOLUTION

Marco Bayas October 31, 2001

Abstract.

The validity of the molecular clock hypothesis is discussed. Molecular clocks depend on both the particular protein considered and the group of species being compared. This means that the neutral theory formulated by Kimura is not generally correct. Deviations from neutrality are explained in base of different points of view such as selectivity.

Introduction.

The molecular clock hypothesis states that the rate of evolution of informational macromolecules is constant through time. In the case of proteins the study of the differences of sequences among different species suggest that evolution takes place at a rate of approximately one substitution in 2.8×10^7 for a polypeptide chain consisting of 100 amino acids[2]. This hypothesis was incorporated in the so-called neutral theory of the molecular evolution formulated by Kimura [4].

The assumption of that molecular evolution is sufficiently regular over time and across lineages is used for testing phylogenetic hypothesis, or for estimating the time of remote evolutionary events. In this article we will discuss the limitations of that assumption.

Neutral Theory.

Kimura's neutrality theory of molecular evolution states that most amino acid substitutions may be of little or no functional consequence and consequently they are not strongly constrained by natural selection [2]. If mutations are neutral with respect to natural selection their frequencies will change only by accidental sampling errors from generation to generation, that is by genetic drift. Rates of replacements will thus be stochastically constant that is they will occur with a constant probability for a given protein. Any such protein would function as a molecular clock: the number of amino acid replacements would be expected to be proportional to the time since the divergence from the common ancestor

If the time of divergence of two modern species is known from fossil records, the proportion of amino acid differences between homologous protein sequences can be used to calculate the rate of evolution. As an example, let us consider the evolution of α-

hemoglobin. The differences in amino acid sequences between some species are tabulated in the next table [4].

	Shark	**Newt**	**Chicken**	**Echidna**	**Dog**	**Human**
Shark	0	0.614	0.597	0.604	0.568	0.532
Newt	0.952	0	0.447	0.504	0.461	0.44
Chicken	0.909	0.592	0	0.34	0.312	0.248
Echidna	0.926	0.701	0.416	0	0.298	0.262
Dog	0.839	0.618	0.374	0.354	0	0.163
Human	0.759	0.58	0.285	0.304	0.178	0
<K>	**0.877**	**0.623**	**0.358**	**0.329**	**0.178**	
Time (My)	**450**	**360**	**290**	**225**	**80**	

The values above the diagonal are the observed proportion of amino acid differences (D) between the α-hemoglobin sequences in the species. The values below the diagonal are the expected amino acid differences per site (K = - ln(1-D) [4]). The values in boldface are the average values of K and the time of divergence.

If the rate of evolution is constant, the relation between K and time should be linear. In the case of the data showed for α-hemoglobin R^2 for linear regression is 0.885 and the rate of evolution estimated is 1.7×10^{-9} replacements per site and per year. It is important to mention that the estimation will be better if more species are compared.

According to the neutrality theory, evolution at the molecular level consists for the most part of the gradual replacements in gene sequences given as a result an allele that is functionally equivalent to the first. The equivalence of a replacement depends on the natural selection constrains on the protein coded by a mutated gene. Because, there are many proteins that differ in their functional constrains, thus there are many molecular clocks that ticking at different rates. For example the rate of evolution of Cytochrome c is smaller than the one for hemoglobins [4].

Deviations from neutrality

Some deviations from neutral behavior have been detected. For example, the rate of evolution calculated for a given protein depends on the particular set of species considered. This is the case of glycerol-3-phosphate dehydrogenase (GPDH) and superoxidase bismutase (SOD) [1]. In fluit flies, GPDH evolves at a rate of 1.1×10^{-10} amino acid replacements per site per year (rpspy) when Drosophila species are compared (time of divergence: 55 million years) but a much faster rate of 4.5×10^{-10} replacements per site per year when comparisons are made between mammals (time of divergence~70 My).The rate of SOD evolution is very fast between drosophila species (16.2×10^{-10} rpspy) and it becomes much slower when all animal species are considered (5.3×10^{-10}).

Another way to notice deviations form neutrality is taking into account that he clock predicted by the neutrality theory behaves as a Poisson process, so that the ratio, R,

of the variance to the mean is expected to be 1, which can readily be empirically tested. The results of many such tests have shown that R is almost universally greater than 1 [1]. This means that in these cases the rate of the molecular substitution is not constant through time.

Several modifications to the neutral theory have been proposed seeking to account for the excess variance of the molecular clock. It has been propose, for example, that most protein evolution involves slightly deleterious replacements rather than strictly neutral ones; or that the effectiveness of the error-correcting polymerases varies among organisms, so that mutation rates change.

Another supplementary hypothesis invokes a generation time effect. For example, the rate of evolution protein evolution in primates is slower than in rodents. According to the generation effect hypothesis, these rate differences could be explained by assuming that the rate of evolution depends on the rate of reproduction, which is several times greater for the short-generation rodents than for the long primates. Then, the larger the number of replication cycles, the greater the number of mutational errors that will occur.

DISCUSSION.

The rate of molecular evolution is not constant. Different proteins evolve a different rates depending on basically the functional restrictions imposed by natural selection. Even, the rate of evolution of a given protein could change in time. These considerations should be taken into account at the moment of using molecular clocks.

The evolution of the genetic structure of populations is driven by several “forces”, including mutation, migration, selection and random drift[4]. The degree of influence of these different factors will determine the characteristics of the evolution of the genetic structure of the different species. It seems that the random drift is the most important determinant for molecular evolution, considering the high applicability of the neutral theory.

References

1. Ayala, F. 1997.Vagaries of the molecular clock. Proc. Natl. Acad. Sci. 94:7776-7783.

2. Kimura, M. 1968. Evolutionary Rate at the Molecular Level. Nature. 217:624-626.

3. Kreitman, M. 2000. Methods to detect Selection in Populations with applications to the humans. Annu. Rev. Genomics Hum. Genet. 01:539-59

4. Hartl and Clark. 1997. Principles of population genetics. Sinauer Associates. Inc. Publishers. Sunderland. Massachusetts.