Molecular Evolution: Rates

Advanced article

Article Contents

- Tempo and Mode of Molecular Evolution
- Studying Patterns of Rate Variation between Lineages
- Patterns of Rate Variation and the Molecular Clock

Online posting date: 15th July 2008

Lindell Bromham, *Centre for Macroevolution and Macroecology, School of Botany and Zoology, Australian National University, Canberra, Australia*

The pace of molecular change is generally more even than morphological evolution, yet there is clearly variation in molecular evolution rate between biological lineages. The causes of this rate variation are still largely unknown, but may be influenced by species characteristics, population structure and evolutionary processes.

Tempo and Mode of Molecular Evolution

Tempo and mode of phenotypic evolution varies across taxa and over time. Bradytelic taxa ('living fossils') such as horseshoe crabs, coelacanths and gingkos can change so little that they are virtually indistinguishable from fossil relatives tens or even hundreds of millions of years old. Hypotelic taxa (fast evolvers) can produce new forms in less than a million years, such as the diverse cichlid fish of Lake Victoria. In contrast, evolutionary change at the molecular level is generally considered to accumulate more or less evenly over time. This apparent constancy of molecular change over evolutionary time has been dubbed 'the molecular clock', and has been exploited as a tool for estimating the age of biological lineages. But while the molecular clock holds true in the broad sense – the longer two lineages have evolved separately, the more genetic differences between them – it has become clear that tempo and mode of molecular evolution can also vary markedly among genes, over time and between species. Variation in rate of molecular evolution is most obvious for comparisons between highly divergent lineages – the substitution rate of viruses can be one million times faster than mammals, and mammalian substitution rates are an order of magnitude faster than sharks. More intriguingly, it seems that even closely related species can show consistent differences in rate of molecular evolution. **See also**: Evolution: Tempo and Mode; Molecular Clocks; Molecular Evolution: Introduction

Investigation of the patterns and causes of rate variation is important for both practical and theoretical reasons. Molecular date estimates have revolutionized evolutionary biology, with the promise of an independent timescale for evolution. But the molecular clock is a 'sloppy clock'. Substitutions do not tick steadily like a conventional clock but occur at random intervals. This sloppiness alone is enough to put limits on the precision of molecular date estimates, giving them wider confidence intervals than is often acknowledged. Furthermore, rate of molecular evolution varies between lineages, a calibration rate established for one lineage may not be an accurate reflection of the substitution rate in another lineage, which could lead to substantial errors in molecular date estimates. The study of patterns and causes of rate variation is essential to critical evaluation of the accuracy and precision of molecular date estimates, particularly where molecular dates are in dramatic contrast to dates obtained from other sources such as palaeobiology or biogeography. **See also**: Geological Time: Dating Techniques; Mutation Rates: Data

The study of rate variation also provides a window on the mechanisms of molecular evolution, and plays a critical role in assessing molecular evolutionary theories. Constancy of rates of molecular evolution was one of the key observations on which the neutral theory was based, and is considered by many to be the most compelling argument for the dominant role of mutation and drift in molecular evolution. However, Gillespie (1991) used a detailed examination of patterns of substitution rate in mammalian protein-coding genes to suggest that while 'silent' (synonymous) changes could be described by a nearly neutral (weak selection) model, protein evolution (nonsynonymous substitutions that lead to amino acid changes) did not fit a neutral model. He partitioned the variation in rate of molecular evolution into lineage effects (lineages can have consistently different substitution rates) and residual effects (substitutions do not accumulate evenly over time but tend to occur in bursts). Substitutions may be clustered in time for a number of reasons. For example, evolution of a new function may prompt a series of changes to the amino acid sequence in the active site of a protein, or a substitution of one amino acid may require a series of compensatory changes to maintain protein structure. These microevolutionary effects on rate of molecular evolution could make any given gene evolve faster in one lineage than another. This article is concerned with the 'macroevolutionary' effects that create consistent genome-wide differences in rate between lineages. **See also**: Drift: Introduction; Microevolution and Macroevolution: Introduction; Molecular Evolution: Nearly Neutral Theory;

ELS subject area: Evolution and Diversity of Life

How to cite:
Bromham, Lindell (July 2008) Molecular Evolution: Rates. In: Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd: Chichester.
DOI: 10.1002/9780470015902.a0001802.pub2

Molecular Evolution: Neutral Theory; Mutations and the Genetic Code

Studying Patterns of Rate Variation between Lineages

There are two broad ways of studying variation in rate of molecular evolution: absolute and relative. Absolute measures are estimates of the amount of molecular change that has occurred in a known time period, and can be directly compared across taxa. Molecular evolution is usually too slow to be directly observed, so absolute rates can only be estimated for lineages with known divergence times, usually taken from the fossil record but occasionally from biogeography (e.g. island ages), sample dates (e.g. HIV sampled throughout the epidemic history) or phylogeny (e.g. calibrating cospeciating parasites from host divergence times). Absolute estimates are usually only practical for lineages with a reliable and continuous fossil record, or fast evolving viruses that have been serially sampled. In contrast, relative rate estimates can be made for all taxa, without calibration rates, by making comparisons between the rate of molecular evolution in related lineages. **See also**: Fossils in Phylogenetic Reconstruction; Molecular Evolution: Patterns and Rates

The simplest way to compare rates of molecular evolution between lineages is the relative rates test, which compares the distance between each of a pair of taxa and an outgroup to determine the relative amount of change in each lineage since their last common ancestor (**Figure 1**). The relative rates test can provide a test of the molecular clock using estimates of the expected variation in substitution rate to identify significant departures from rate constancy. Other tests of rate constancy (sometimes called 'clock tests') include the Tajima triplet test (compares the number of sites at which the amino acid or nucleotide state is shared by the outgroup and only one of the two ingroups) and the likelihood ratio test (clock-like evolution is rejected if a multiple rate model gives a significantly better fit to the data than a constant rate model). It is important to note that all of these clock tests are limited in power for short sequences or for low levels of divergence. This makes the use of clock tests to select rate constant genes for molecular dating problematic, since failure to detect departures from a molecular clock can potentially lead to overestimation of date estimates. But tests for variation in rate of molecular evolution do provide a simple and effective tool for identifying differences in rate between lineages. **See also**: Molecular Clock: Testing; Molecular Phylogeny Reconstruction

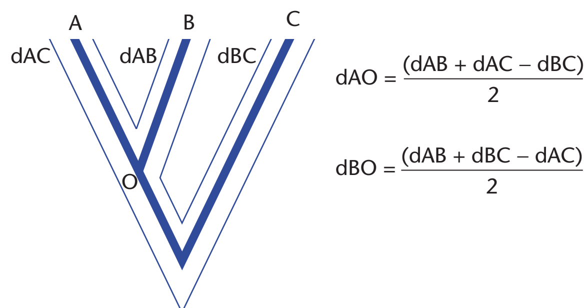

Figure 1 Relative rates test: the genetic distances between each of a pair of taxa (A and B) and an outgroup (C) can be used to compare the amount of evolution along lineages A and B since their last common ancestor (O).

The discussion of causes of variation in rate of molecular evolution between lineages has been largely anecdotal to date, with hypotheses based on few divergent taxa, or even on single phylogenies. Consequently, few of the suggested mechanisms have substantial empirical support. Availability of both data and methodology should now allow the study of tempo and mode of molecular evolution to move from hypothesis generation to a phase of description of general trends and broad-scale hypothesis testing. Tests for rate variation can be used to compare the rate of molecular evolution in lineages that differ in some variable of interest, such as generation time or metabolic rate. If a sufficient number of phylogenetically independent pairs are chosen, the relationship between the proposed causal variable and rate of molecular evolution can be tested statistically (see Bromham *et al*., 1996). The following discussion of lineage-specific rate variation is by no means exhaustive, but is intended to illustrate the wide range of factors that could influence molecular evolutionary rates.

Balance between mutation and repair

Point mutations in deoxyribonucleic acid (DNA) sequences (one base replacing another) arise from two sources: imperfectly repaired damage and replication errors. So species can differ in their mutation rate if they differ in the accuracy of DNA replication or the efficiency of damage repair. All organisms must strike a balance between the competing costs of mutation and DNA repair. For example, the net mutation rate can be affected by the relative balance between the 'forward' and 'backward' activities of DNA polymerase. More emphasis on proofreading (exonuclease activity) may result in fewer copy errors, but it will be at the expense of polymerase activity (chain elongation), not only slowing replication but resulting in increased excision of correct bases as well as mispairs. This balance between forward and backward activity can be altered by point mutations, as demonstrated by both mutator and antimutator polymerases (that copy DNA with lower or higher fidelity than wildtype, respectively). Since repair efficiency and replication fidelity may be heritable and naturally variable in populations, selection can drive to a different balance between mutation and repair in different lineages. **See also**: DNA Damage; DNA Repair; Mutation–Selection Balance

The effect of selection on mutation rate is most obvious for single-celled organisms. In an experimental population of bacteria, a strain carrying a mutation with a fitness advantage can rapidly outcompete and replace the wildtype.

 ENCYCLOPEDIA OF LIFE SCIENCES © 2008, John Wiley & Sons, Ltd. www.els.net

Raised mutation rate increases the chance of a fitter mutant arising, so 'mutator' strains (with higher mutation rates than the wildtype) can come to predominate. But high mutation rate carries a cost, as mutator strains suffer decreased growth and the gradual build up of deleterious mutations. Wild-type bacteria seem to have evolved to a compromise position between the costs and benefits of mutation and repair. The evolution of mutation rates in complex eukaryotes is less transparent, due to complications of genomic organization such as diploidy, multiple chromosomes, sexual reproduction and the separation of germline and soma. But they must still find a balance between the costs of mutation and repair. Any increase in repair efficiency must be paid for in resources used or time taken, and at some point the additional costs required to further improve repair efficiency will outweigh the benefits gained. Indeed, DNA repair rates in complex eukaryotes are apparently neither at the maximum achievable level, nor at the minimum tolerable level. There is sufficient natural variation in repair efficiency and replication fidelity to provide the raw material for selection to fine-tune mutation rates. However, it is not yet clear whether the relative costs and benefits of mutation and repair can vary sufficiently between multicellular eukaryote lineages to contribute to the evolution of lineage-specific rates of molecular evolution. **See also**: Fitness; Mutagenesis Mechanisms

Morphological change and adaptation

Many cases of the effect of adaptation on the rate of evolution of specific genes have been documented. A frequently cited example is the evolution of lysozyme in mammals. The usual function of lysozyme is defence against bacteria, but in two lineages that have independently evolved foregut fermentation, artiodactyls and leaf-eating langur monkeys, lysozyme has been co-opted into digestion. In order to function in the chemically challenging environment of the stomach, lysozyme has undergone a suite of amino acid changes, and this has resulted in an accelerated rate of amino acid substitutions in lysozyme gene in both artiodactyls and langur monkeys. So it is clear that the process of adaptation can cause localized areas of the genome to experience higher rates of molecular evolution in some taxa than in others. These bursts of selection on genes might be the cause of the residual variation in the molecular clock, making the tick rate more variable than expected from a purely stochastic process. **See also**: Adaptation and Natural Selection: Overview; Adaptation: Genetics

But while there are examples of taxa or genes that have an elevated rate of molecular evolution due to phenotypic adaptation, evidence for a more general relationship between rates of morphological and molecular rates of evolution is limited. All adaptation must be ultimately derived from change at the molecular level, but any given morphological change may affect only a few nucleotides in one or few genes, representing only a tiny proportion of the genome. Most of the molecular changes observed are likely to be unlinked to adaptive evolution, occurring at nucleotide sites that do not affect protein or ribonucleic acid (RNA) products, or causing molecular changes that have no appreciable affect on fitness. The overall independence of rates of morphological and molecular evolution is supported by the observation that slowly evolving taxa, such as bradytelic 'living fossils', do not have appreciably slower rates of molecular evolution than their hypotelic relatives. Similarly, taxa that show a high level of morphological disparity, such as molluscs and echinoderms, do not appear to accumulate more molecular changes in genes used for phylogenetic analysis than less morphologically diverse clades. **See also**: Molecular Phylogeny Reconstruction; Mutations and New Variation: Overview

Life history correlates of rate variation

One of the most consistent patterns in vertebrate molecular evolution is the negative body size trend in rate of molecular evolution – smaller animals tend to have faster rates of molecular evolution than their larger-bodied relatives (**Figure 2**). This body size effect has been demonstrated for a number of genes (including mitochondrial and nuclear, RNA and protein, introns and exons) as well as DNA–DNA hybridization data and restriction fragment length polymorphism (RFLP) distances. How could body size affect rate of molecular evolution? The two main theories concern life history variables that scale with body size: generation time and metabolic rate.

The generation time effect assumes that species with shorter generation turnover times will undergo more DNA replications per unit time, and will therefore accumulate more replication errors. The effect of number of DNA replications on substitution rate is supported by the observation of 'male-driven evolution'. In mammals, it takes more cell divisions to make sperm than eggs, so genes that spend more time in males ought to accumulate more copy errors, a prediction is borne out by a comparison of sequences on Y-chromosomes (which spend all their time in males), autosomal chromosomes (which spend on average half of their time in males) and X-chromosomes (which spend only a third of their time in males; see Ellegren, 2007). An alternative explanation for the body size effect on rate of molecular evolution in vertebrates is the metabolic rate hypothesis. Small animals tend to have higher mass-specific metabolic rates than their larger relatives. Increased metabolic rate results in a higher concentration of metabolites, such as free oxygen radicals, which are thought to damage DNA. This hypothesis has been supported by the observation that poikilotherms have lower rates of molecular evolution than endotherms of equivalent size or generation length (Martin and Palumbi, 1993). Both the generation time effect (the more you copy, the more mistakes you make, the higher your mutation rate) and the metabolic rate effect (the more you burn, the more damage you do, the higher your mutation rate) provide plausible explanations for the body size trend in molecular evolution rate. Moreover, these life history traits scale together in

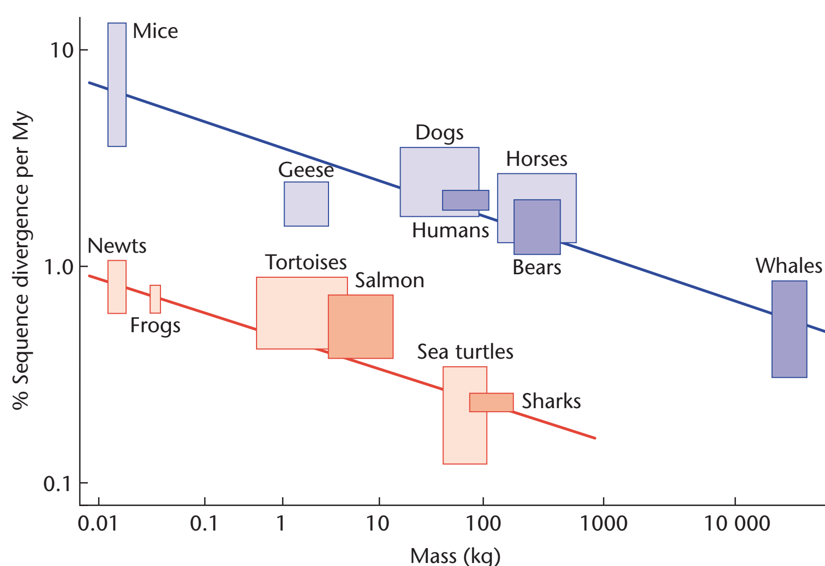

Figure 2 The negative body size effect on rate of molecular evolution in vertebrates. Modified from Martin and Palumbi (1993).

vertebrates – small animals tend to have both a high metabolic rate and a short generation time – so if either is the causal variable it will cause a spurious association with the other. However, studies that have compared these two hypotheses using a relative rates comparative method found no evidence of an effect of metabolic rate beyond its covariation with body size generation time (Lanfear *et al.*, 2007). **See also**: Chromosome Function: Sex Differences; Free Radicals and Other Reactive Species in Disease; Metabolic Turnover; Vertebrate Metabolic Variation; Vertebrate Metabolism

However, the case for a generation time effect in vertebrate molecular evolution is far from clear-cut. Firstly, the number of DNA replications per organism generation varies between species, and it seems likely that it also scales with body size, partially ameliorating the potential for a generation time effect. Secondly, while mitochondria must replicate at cell division, they also replicate throughout the lifetime of the cell at a frequency partly determined by the role and activity of the cell. This suggests that the generation time effect ought to be less obvious for mitochondrial genes, and yet they show as strong a correlation as nuclear genes. Thirdly, generation time clearly does not provide a global explanation of lineage-specific rate variation – for example, the range of generation times in sharks overlaps with that of primates, yet shark rates are estimated to be an order of magnitude lower than primate rates. It may be that these complicating factors dampen but not negate the generation time effect. But it is also possible that generation time is a proxy variable for the true causal factor. **See also**: Molecular Evolution: Rates

Speciation and population processes

The relationship between speciation and genetic change has been debated for decades, but a link between speciation rate and rate of molecular evolution has recently received some empirical support. Large scale molecular phylogenies of flowering plants have allowed comparisons of clade size (number of species) and molecular branch length for close to one hundred sister groups (Barraclough and Savolainen, 2001). These studies found a significant association between net speciation rate and substitution rate for three genes. A study that took a different approach – asking whether the number of nodes (divergence points) in a phylogeny was associated with the net amount of molecular change inferred – also found an association between diversification rate and rate of molecular evolution for a broad range of taxa (Webster *et al.*, 2003). Further studies will be needed to determine the generality and cause of the association between clade size and rate of molecular evolution. It is possible that lineages with a faster rate of molecular evolution are somehow predisposed to speciation, perhaps because rapid molecular evolution leads to faster reproductive isolation by evolution of genomic incompatibility. Alternatively, the process of speciation may influence the rate of molecular evolution. Since the relationship seems largely due to synonymous changes, it is unlikely to be a direct effect of selection on these loci during speciation. One possible connection between speciation rate and rate of molecular evolution is population size. Rapid speciation could result in repeated population subdivision (as incipient species become reproductively isolated from each other), potentially creating lower average population size in rapidly radiating lineages. Population size affects substitution rate by altering the relative power of selection and drift. Allele frequencies in small populations are greatly affected by random sampling, so can vary from generation to generation by chance. This disrupts the power of selection in small populations, so that the frequency of alleles of

small selective effect (nearly neutral mutations) will be governed by drift rather than selection. So while the fixation of advantageous alleles is faster in large populations, or populations undergoing exponential growth, the rate of fixation of neutral and nearly neutral mutations will be higher in small populations. **See also**: Population Genetics: Overview; Speciation: Genetics; Speciation: Introduction

Changes in population size could affect genome-wide substitution rates, but most population size fluctuations occur on a timescale that is so much shorter than the frequency of substitutions that average substitution rate may be unaffected (Gillespie, 1991). Only factors that have a long-term and consistent effect on population size, such as life history, ecological specialization or biogeography, are likely to contribute significantly to lineage-specific substitution rates. Another life history trait that influences effective population size over evolutionary time is mating system. For the same number of individuals, a randomly mating population, where all individuals have roughly equal chances of passing on their genes, has a higher effective population size than a population with a hierarchical or territorial social structure where only a few dominant males have the chance to mate each year. **See also**: Reproductive Strategies

Patterns of Rate Variation and the Molecular Clock

Recognition of variation in rate of molecular evolution need not lead to an outright rejection of the molecular clock, which remains a fundamentally important tool in evolutionary biology. But it does beg caution, particularly in cases where the reliability of molecular date estimates has been called into question. For example, molecular dates for the rat–mouse divergence approximately 40 million years ago (Mya) are surprisingly older than fossil-based estimates of the split at approximately 16 Mya. The rate of molecular evolution in rats and mice is much faster than most other mammals, so any calibration rate estimated from a different mammalian lineage (or averaged over the whole phylogeny) will be too slow, causing overestimation of the amount of time needed to produce the observed molecular divergence between rats and mice. Molecular dates are likely to be overestimated whenever there is consistent variation in rate of molecular evolution between lineages. If sequences are assumed to be rate constant such that the distance between them is halved to calculate time since divergence, then a faster rate in any lineage will make the divergences appear deeper than they really are. This is particularly a problem if clock tests are ineffective at detecting rate variation in sequences used for molecular dating. **See also**: Molecular Evolution; Molecular Evolution: Patterns and Rates

Investigation of lineage-specific rate variation and its effects on molecular date estimates will be critical to resolving a controversy that has generated a great deal of heated debate between palaeontologists and molecular biologists. For three major 'explosive' radiations inferred from the fossil record, the Cambrian explosion of animal phyla and the early Tertiary radiation of both birds and mammals, molecular dates are nearly twice as old as the fossil dates of divergence. The molecular dates imply that these explosive radiations are an artefact of the fossil record, and birds, mammals and metazoans have a substantial hidden history. But before dismissing the palaeontological evidence, the accuracy and precision of these molecular dates should be critically examined, particularly in light of possible causes of lineage-specific rate variation. For example, if both speciation rate and rate of morphological change are associated with high rates of molecular evolution, then we should expect rates to be accelerated in an explosive radiation. Concerted temporal variation in rates cannot be detected by conventional clock tests, so fast early rates in all lineages could cause the date of the radiation to be overestimated. Similarly, the mammalian radiation was characterized by a rapid increase in average body size in virtually all orders, so if small body size is associated with fast rates, the early mammal lineages might have had faster rates. Once again, this is unlikely to be detected and would result in overestimation of the date of the mammalian radiation. These examples are highly speculative and are intended only as illustrations of the potential effects of lineage-specific rate variation on interpretation of molecular data. We cannot know what impact rate variation has on our interpretation of DNA data until we have a better understanding of the generality and causality of patterns of rate variation. **See also**: Cambrian Radiation; Speciation and the Fossil Record

References

Barraclough TG and Savolainen V (2001) Evolutionary rates and species diversity in flowering plants. Evolutionary rates and species diversity in flowering plants. *Evolution* **55**: 677–683.

Bromham L, Rambaut A and Harvey PH (1996) Determinants of rate variation in mammalian DNA sequence evolution. *Journal of Molecular Evolution* **43**: 610–621.

Ellegren H (2007) Characteristics, causes and evolutionary consequences of male-biased mutation. *Proceedings of the Royal Society of London. Series B* **274**: 1–10.

Gillespie JH (1991) *The Causes of Molecular Evolution*. Oxford: Oxford University Press.

Lanfear R, Thomas JA, Welch JJ, Brey T and Bromham L (2007) Metabolic rate does not correct the molecular clock. *Proceedings of the National Academy of Sciences of the USA* **104**: 15388–15393.

Martin AP and Palumbi SR (1993) Body size, metabolic rate, generation time and the molecular clock. *Proceedings of the National Academy of Sciences of the USA* **90**: 4087–4091.

Webster AJ, Payne RJH, Pagel M (2003) Molecular phylogenies link rates of evolution and speciation. *Science* **301**: 478.

Further Reading

Bromham L (2008) *Reading the Story in DNA: A Beginner's Guide to Molecular Evolution*. Oxford: Oxford University Press.

Kimura M (1983) *The Neutral Theory of Molecular Evolution*. Cambridge, England: Cambridge University Press.

Lynch M (2007) *The Origins of Genome Architecture*. Sunderland, MA: Sinauer Associates.

Maynard S (1998) *Evolutionary Genetics*, 2nd edn. Oxford, UK: Oxford University Press.

Page RDM and Holmes EC (1998) *Molecular Evolution: A Phylogenetic Approach*. Oxford: Blackwell Science.

 ENCYCLOPEDIA OF LIFE SCIENCES © 2008, John Wiley & Sons, Ltd. www.els.net