

Molecular Evolution: Patterns and Rates

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

The pace of deoxyribonucleic acid (DNA) evolution varies across the genome and between species, determined by the balance between mutation, selection and drift.

Introduction

Patterns of molecular evolution are complex and highly variable. This complexity is evident in a comparison of rates of change across the genome and between biological lineages. Different species can have dramatically different rates of molecular evolution. Human immunodeficiency virus (HIV) has a rate of molecular evolution that is around a million times faster than that of mammals. Within mammals, rats and mice have an average rate of molecular evolution that is several times faster than that of humans, so any sequence in a mouse is likely to accumulate more changes in a given time period than the equivalent gene in humans. Rates can vary between the different genomes within a single species: the rate of evolution of mitochondrial genes is consistently higher than that of nuclear genes. Within any genome, different genes have dramatically different rates of evolution. The envelope gene in HIV, critical in the interaction between the virus and the host genome, evolves many times faster than the rest of the HIV genome. Different sites within a gene can also evolve at different rates. A comparison of deoxyribonucleic acid (DNA) polymerase from the five kingdoms of life shows that the amino acid sequence of the catalytic site has remained virtually unchanged for billions of years, while the structural parts of the proteins have evolved. These differences in the rate of molecular evolution are due to differences in the processes of mutation and substitution. **See also:** [Evolution: Tempo and Mode](#); [Human Immunodeficiency Viruses \(HIV\)](#); [Molecular Evolution: Introduction](#); [Molecular Evolution: Rates](#)

The term mutation can refer to any process that changes the genetic information in the genome, including DNA insertions, deletions and rearrangements, but here we will consider only point mutations, where one base in the DNA

Introductory article

Article Contents

- Introduction
- Mutation Rates and the Rate of Molecular Evolution
- Selection and Drift Affect Substitution Rate
- Summary

Online posting date: 15th July 2008

sequence is changed for another. Mutations occur spontaneously in every generation, but not all mutations become permanent features of the genome. Some mutations will be lost from the population. A new mutation will be less likely to be passed on to the next generation if it reduces the probability that the individual that carries it will survive and reproduce. Or, given that a new mutation starts at very low frequency in the population, it may simply be lost by chance if its carrier dies or fails to breed. Conversely, a mutation can increase in frequency, either through the action of selection or by chance. The process of new mutations increasing in frequency to the point where they replace the predominant (wild-type) allele in the gene pool is called substitution – one nucleotide in the DNA sequence is substituted for another. **See also:** [Mutations and the Genetic Code](#)

Mutation Rates and the Rate of Molecular Evolution

Mutation rate – the frequency of changes to the DNA nucleotide sequence – can vary between species and across the genome, and these differences contribute to the variation in the patterns and rates of molecular evolution. The two major sources of point mutations are replication errors (mistakes made when copying DNA) and DNA damage. **See also:** [Mutation Rates: Data](#)

DNA replication and copy errors

DNA replication is remarkably accurate. The DNA copying machinery in eukaryotes usually makes a mistake less than once in every million nucleotides. This incredible copy fidelity is achieved by a sophisticated error-checking system involving base selection, proofreading and postreplication repair. But the efficiency of these error-correction pathways varies across the genome and between taxa, creating differences in mutation rate which could ultimately influence the rate of molecular evolution. HIV uses a relatively error-prone replication system. Reverse transcriptase (ribonucleic acid (RNA)-directed DNA polymerase; EC 2.7.7.49), an enzyme that creates a DNA copy of the HIV genome, has no proofreading and consequently its error rate in orders of magnitude is higher than mammalian

ELS subject area: Evolution and Diversity of Life

How to cite:

Bromham, Lindell (July 2008) Molecular Evolution: Patterns and Rates. In: Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd: Chichester.
DOI: 10.1002/9780470015902.a0001799.pub3

rates. Within mammalian cells, there are many different polymerases that copy DNA with varying degrees of fidelity. Mitochondrial DNA is copied by polymerase gamma, which has a higher error rate than other mammalian polymerases, and this might contribute to the higher mutation rate in mitochondrial genes than in nuclear genes. **See also:** [DNA Polymerase Fidelity Mechanisms](#); [Retroviral Replication](#)

Since mistakes in DNA copying contribute to mutation rate, we should expect that the more times DNA is copied, the more errors would accumulate. This is known as the generation time effect. The generation turnover time of HIV can be as little as a day which, combined with a high-replication error rate, drives the rapid accumulation of new mutations that allows HIV to evolve within its human host. Mice have a generation turnover time of around a month. Mice can go through 150 generations for every human generation, so mouse DNA must be copied more often than human DNA, and therefore has more chances to accumulate copy errors. Frequency of DNA replication might also contribute to the difference in the rate of molecular evolution between mitochondrial and nuclear genes. Nuclear genes are replicated at cell division, but mitochondrial genes may be replicated many times throughout the lifetime of a cell. So if mitochondrial genes have more DNA replications per unit time, they ought to accumulate more copy errors than nuclear genes. **See also:** [Mitochondria: Structure and Role in Respiration](#)

DNA damage and repair

Mutations can also occur by damage to the DNA. Such damage occurs spontaneously throughout the lifetime of the cell, but the rate of DNA damage can be greatly increased by mutagens. Cells have an array of repair systems which reduce the incidence of mutations resulting from DNA damage. Variation in mutation rate between species or across the genome could arise from differences in the amount of damage that occurs or differences in the efficiency of repairing damage. For example, ultraviolet (UV) light is a mutagen that causes harmful lesions in DNA. Eukaryotic cells have an excision repair system that greatly reduces the incidence of mutation by UV light by detecting these lesions, cutting out the affected stretch of DNA and replacing it. Deficiencies in this excision repair system can cause a greatly elevated incidence of the mutation rate as a result of reduced ability to repair UV-related DNA damage (which can result in a higher incidence of skin cancer). So the efficiency of DNA repair systems is critical in determining the mutation rate resulting from DNA damage. Repair efficiency can vary between species. For example, rodents appear to lack some of the excision repair pathways found in primates, and this could contribute to the higher rate of molecular evolution in rats and mice than in humans. Repair efficiency also seems to vary across the genome, with transcribed genes receiving heightened repair effort. **See also:** [DNA Damage](#); [DNA Repair](#);

Environmental Carcinogens and Mutagens; Mutagenesis Mechanisms; Nucleotide Excision Repair in Eukaryotes

Species could also differ in mutation rate if they suffer different amounts of DNA damage. One source of DNA damage that could differ between species is the effect of cellular metabolism, which produces potentially mutagenic by-products such as free oxygen radicals. It has been suggested that high metabolic rate raises the concentration of free oxygen radicals in cells, increasing the rate of DNA damage, and that this should be reflected in higher rate of molecular evolution. This hypothesis is consistent with the observation that ectotherms ('cold-blooded' animals such as fish and reptiles) have slower rates of molecular evolution than endotherms ('warm-blooded' animals such as birds and mammals). The higher rate of molecular evolution in smaller-bodied species could be due to their higher mass-specific metabolic rate inflating the mutation rate, though there is currently no evidence that metabolic rate can explain much of the variation in rates between animal lineages. The metabolic rate hypothesis could also explain the high rate of molecular evolution in mitochondrial genes. Since mitochondria are the site of oxidative metabolism, they are expected to have a higher concentration of free oxygen radicals and other by-products of metabolism, and so might suffer the greatest effect of DNA-damaging metabolites. **See also:** [Ecological Consequences of Body Size](#)

Selection and Drift Affect Substitution Rate

Selection at the molecular level

Even if the mutation rate is uniform, rates and patterns of molecular evolution can vary across different sites, genes, genomes or lineages due to factors influencing substitution rate. For any given mutation, the chance of going to fixation (replacing the wild-type in the population) is dependent on the properties of both the mutation itself (its selection coefficient) and the population in which it arises (effective population size). These two things determine the balance between the two forces that shape substitution rates: selection and drift. **See also:** [Selection: Units and Levels](#)

Advantageous mutations, by definition, are mutations with an elevated chance of being passed on to the next generation. Selection increases the proportion of individuals in the population carrying a given advantageous mutation. The relative advantage or disadvantage of a mutation over the alternative alleles in the population is expressed as a selection coefficient, which reflects the chance of whether the mutation will be passed to the next generation. A mutation that greatly increases the chance of survival and reproduction will become fixed faster than one with a relatively minor advantage. For example, positive selection drives a high substitution rate in the HIV envelope

gene. Any change that allows it to go unrecognized by the host immune system will have a huge selective advantage.

See also: [Selection: Frequency-dependent](#)

Selection also removes harmful mutations from the population. By definition, a deleterious mutation is one that reduces its carrier's chance of survival and reproduction, so there should be fewer copies of that mutation passed on to each subsequent generation. Some mutations are so disruptive that they are lethal, and cannot persist even for a single generation. For example, the amino acid sequence of the catalytic site of DNA polymerase is conserved by negative selection removing mutations that disrupt the functioning of this essential enzyme. The effect of selection on reducing the proportion of individuals carrying deleterious mutations is a little more complicated in organisms with diploidy (two complete copies of the genome). Deleterious mutations can persist in the population if they are recessive (only harmful when homozygous), due to heterozygote advantage (having one copy of the allele is advantageous, but two copies is disadvantageous), or through mutation–selection balance (produced by mutation as fast as selection can remove them). **See also:** [Functional Constraint and Molecular Evolution](#); [Heterozygous Advantage](#); [Mutation–Selection Balance](#)

Neutral mutations and genetic drift

Some mutations are neither advantageous nor deleterious. Mutations that do not affect fitness are commonly referred to as neutral mutations. Some mutations are neutral because they occur in apparently functionless DNA, such as redundant copies of genes that are no longer transcribed (pseudogenes). Mutations within functional genes might also be neutral if they are 'silent' – the DNA sequence is altered but still codes for the same amino acid sequence. The third bases of many codons can be changed without altering the amino acid specified, and these synonymous changes are likely to have little or no effect on fitness. Even a mutation that alters the amino acid sequence of a protein may be neutral if the structure and function of the resulting protein are unaffected. What is the fate of these neutral mutations? They will be neither promoted nor removed by selection, but they may still disappear from the population or go to fixation in the gene pool simply by chance, a process known as genetic drift. **See also:** [Drift: Introduction](#); [Molecular Evolution: Neutral Theory](#)

More progeny are born than will survive and reproduce. Not all individuals in a population will breed, and in the case of diploid organisms, not all alleles from a given individual will end up in successful offspring. So the gene pool of each new generation is a subsample of the gene pool of the previous generation. We have seen that some properties of mutations can increase or decrease their chance of inclusion in the next generation. But since neutral mutations are neither promoted nor expunged by selection, they will simply be randomly sampled from the gene pool in each generation. If a population is very large, the effect of the random sampling of subsequent generations will not appreciably alter the frequency of alleles. The smaller the

population, the greater the effect of random sampling on allele frequencies, and the more chance that the 'random walk' of frequencies will lead to fixation of neutral mutations (**Figure 1**). So in small populations, drift can override selection for mutations of small effect (a mutation will be driven by drift rather than selection if its selection coefficient is less than the reciprocal of twice the effective population size). **See also:** [Population Genetics: Overview](#)

Selection and drift shape rates of molecular evolution

Selection drives the fixation of advantageous mutations and the removal of deleterious mutations. The frequency of neutral mutations will fluctuate randomly, and they will be occasionally fixed or removed by genetic drift. The debate between neutralist and selectionist schools of thought concerns the relative proportions of different kinds of mutations – the distribution of selection coefficients (**Figure 2**). All models recognize that most random changes to the genome are deleterious, because they will tend to disrupt highly organized genetic information. It is the relative proportions of mutations that are advantageous, neutral or nearly neutral that is debated. The selectionist school considers that few changes could be entirely without effect on fitness, so most changes in the genome are ultimately driven by selection. The neutralist school suggests that many mutations are functionally neutral, having no appreciable effect on fitness, so the frequency of many alleles in the population will be determined by mutation and drift. And the nearly neutral theory holds that many mutations are of relatively small effect on fitness. In a large population, selection will be sufficiently powerful to act upon these nearly neutral mutations, but they will evolve by drift in smaller populations. The nearly neutral school illustrates how the balance between selection and drift is determined not only by the selection coefficient of the mutation but also by the population size. Selection is most effective in large populations where the random sampling effects on gene frequencies will be of limited effect. But in smaller populations, drift can override selection for mutations with small selection coefficients. **See also:** [Fitness](#); [Molecular Evolution](#); [Molecular Evolution: Nearly Neutral Theory](#); [Neutrality and Selection in Molecular Evolution: Statistical Tests](#)

The debate between selectionist and neutralist schools of thought has resulted in a search for patterns of molecular evolution consistent with either model. A study of rates of molecular evolution between lineages and across the genome is essential to a comparison of these models. The observation of a molecular clock (constant rate of molecular change) is more easily explained by a stochastic process, dominated by random mutation and drift. The fast rate of molecular change in functionally less important parts of genes is also consistent with a neutral model: if selection was the main driving force in molecular evolution then regions not under strong selection would be expected to change more slowly. However, the footprints of selection on rates of molecular change are also evident. Many

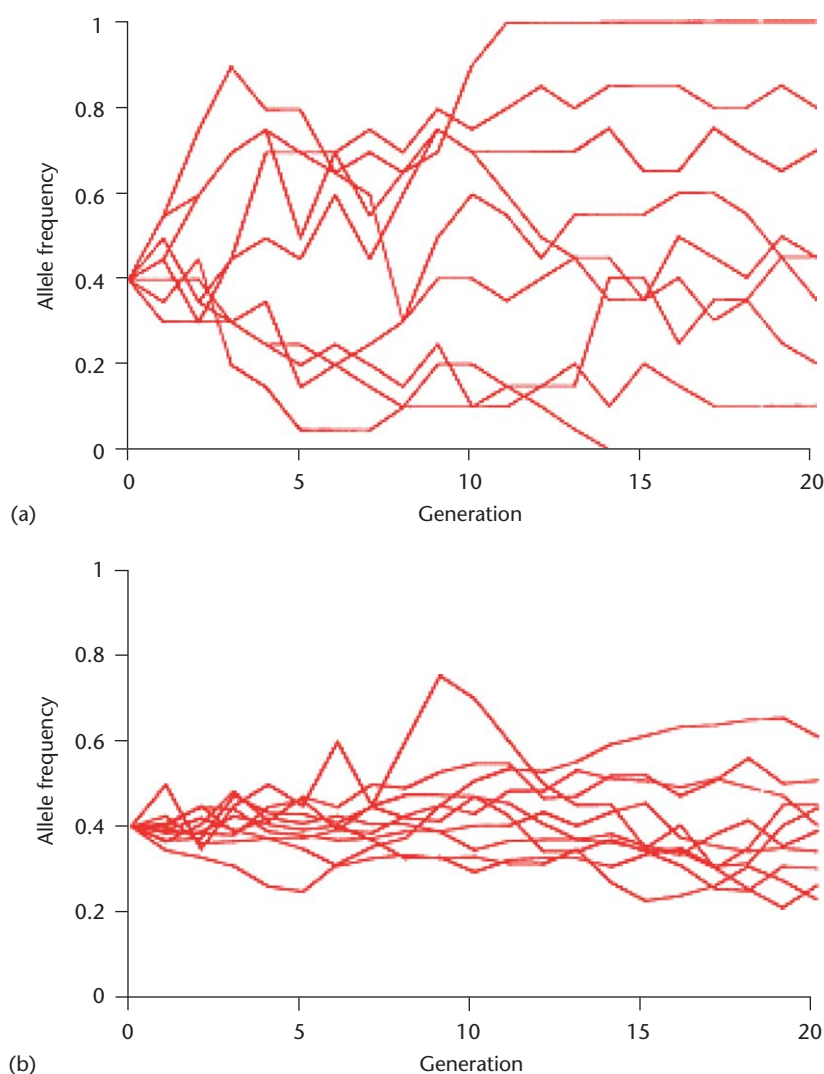


Figure 1 Genetic drift causes random fluctuation in allele frequencies. This can be demonstrated by computer simulations of allele frequencies: (a) In small populations ($N=10$), this 'random walk' in gene frequencies is more likely to result in fixation of neutral or nearly neutral mutations. (b) In large populations ($N=100$), random sampling has less effect on gene frequencies, so the rate of fixation of alleles by drift will be lower.

examples of genes that show evidence of positive selection for amino acid changes have been described, particularly for genes adopting new functions, or associated with host–parasite interactions. **See also:** [Molecular Clock: Testing;](#) [Molecular Clocks](#)

Summary

The observed patterns and rates of molecular evolution are the product of the interaction of a number of processes, a point that is sometimes obscured by the adversarial debates over competing hypotheses. The mutation rate, determined by the balance between mutation and repair, can vary across the genome and between lineages by two means: (1) the rate of DNA replication errors will be higher if more

mistakes are made, fewer are corrected, or if DNA is copied more often and (2) mutation resulting from DNA damage will be increased if more damage is sustained (e.g. high concentration of mutagens), or if damage repair is less efficient. The substitution rate is determined by two factors: (1) the nature of the mutation, whether advantageous, deleterious, neutral or nearly neutral and (2) the effective population size, which determines the power of genetic drift to disrupt selection.

These factors interact to create complex patterns of rates of molecular evolution across the genome and between species. This can be demonstrated by a consideration of the examples of rate variation discussed throughout this article.

- HIV has a faster rate of molecular evolution than mammals because of lower copy fidelity, faster generation turnover and strong positive selection.

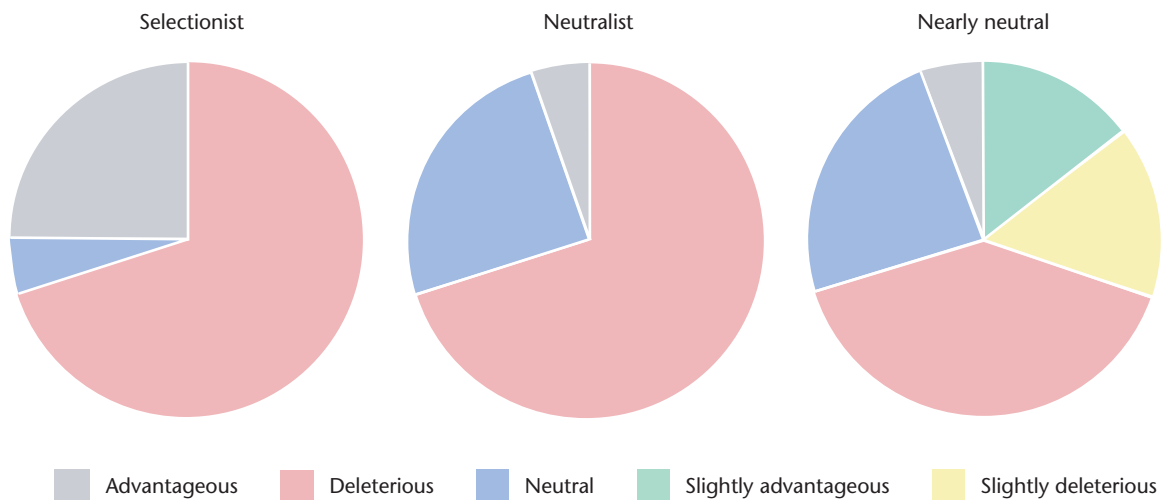


Figure 2 The selection, neutral and nearly neutral schools of thought differ in their predictions of the proportion of mutations that will be advantageous, deleterious, neutral or nearly neutral. Nearly neutral mutations have small selection coefficients, being only slightly deleterious or advantageous.

- Mice have a faster rate of molecular evolution than humans, possibly as a result of more DNA replications per unit time (generation time effect), higher metabolic rate increasing DNA damage (metabolic rate hypothesis) and less efficient DNA repair.
- Mitochondrial genes evolve faster than nuclear genes, potentially due to accumulation of copy errors from more replications per unit time, greater risk of mutation from oxidative damage and less efficient DNA repair.
- The envelope gene of HIV has a high rate of molecular evolution because it is under strong positive selection to produce novel variants; the catalytic site of polymerase evolves slowly because it is under strong negative selection to remove variants; pseudogenes evolve rapidly because they have no effect on fitness, so all mutations have an equal chance of fixation by drift.
- Within protein-coding genes, synonymous (silent) changes evolve faster than nonsynonymous (replacement) changes when the gene is under negative selection, but nonsynonymous changes can outnumber synonymous when the gene is under strong positive selection. **See also:** [Molecular Evolution: Rates](#)

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: Cambridge University Press.
- Page RDM and Holmes EC (1998) *Molecular Evolution: A Phylogenetic Approach*. Oxford: Blackwell Science.
- Ridley M (1996) *Evolution*, 2nd edn. Oxford: Blackwell Science.