APPLICATIONS OF NEXT-GENERATION SEQUENCING

OPINION

Genetic drift, selection and the evolution of the mutation rate

Michael Lynch, Matthew S. Ackerman, Jean-Francois Gout, Hongan Long, Way Sung, W. Kelley Thomas and Patricia L. Foster

Abstract | As one of the few cellular traits that can be quantified across the tree of life, DNA-replication fidelity provides an excellent platform for understanding fundamental evolutionary processes. Furthermore, because mutation is the ultimate source of all genetic variation, clarifying why mutation rates vary is crucial for understanding all areas of biology. A potentially revealing hypothesis for mutation-rate evolution is that natural selection primarily operates to improve replication fidelity, with the ultimate limits to what can be achieved set by the power of random genetic drift. This drift-barrier hypothesis is consistent with comparative measures of mutation rates, provides a simple explanation for the existence of error-prone polymerases and yields a formal counter-argument to the view that selection fine-tunes gene-specific mutation rates.

As mutation affects essentially every aspect of biology, the development of a unifying theory for mutation-rate evolution is highly desirable. There is much to be explained. For although the base-substitution mutation rate (u) in all organisms is low ($<10^{-7}$ mutations per nucleotide site per generation), the rates in some species are more than 1,000-fold below this level. This large range of variation implies that the accuracy of DNA replication and repair in most, if not all, species is less than what is possible at the biochemical level.

It has been argued that mutation rates, even at the single-gene level, have been fine-tuned by natural selection to maximize long-term survival and evolvability[1–4], yet there is no direct empirical or theoretical evidence that this is generally the case. If such adaptive mutation-rate arguments are valid, they will need to explain why the evolved mutation rate in microorganisms is 100- to 1,000-fold lower than that in vertebrates. Moreover, although stress-induced mutagenesis in microorganisms can sometimes provide a transient mechanism for generating an adaptive genotype in an extreme environment[5], this need not imply that the error-prone nature of the polymerases invoked during such times has been promoted by natural selection[6–8].

As whole-genome sequencing (WGS) has led to a rapid proliferation of data on mutation rates in a wide array of phylogenetic lineages, there is a need to evaluate how this information can be integrated into a general evolutionary framework. We start with a brief overview of the theory of mutation-rate evolution, followed by a comparison of the existing data with theoretical expectations. Confronted with difficulties in reconciling observations with adaptive mutation-rate hypotheses, at both the whole-genome and single-gene levels, we argue that phylogenetic variation in mutation rates reflects underlying differences in the efficiency of selection to improve replication fidelity, which in turn results from variation in the power of random genetic drift. There is still considerable room for work on the cellular determinants of replication fidelity and how these vary across phylogenetic lineages, but achieving evolutionary understanding in this area is unlikely to be served by uncritical adherence to the idea that every aspect of genome stability is refined by adaptive processes.

Selection, drift and mutation rate

A formal theoretical framework for understanding mutation-rate evolution was first presented by Kimura[9]. Noting that the vast majority of mutations are deleterious, he proposed that mutator alleles are indirectly selected against through associations with the detrimental alleles that they generate elsewhere in the genome. Under this view, a newly arisen mutator allele progressively acquires an excess equilibrium mutation load that is defined by the opposing forces of input (mutation pressure) and removal (selection and recombination). Here, we broadly define a mutator (and an antimutator) as any genomically encoded modifier of the mutation rate; for example, a variant of a DNA polymerase, a DNA repair protein or even a gene product that alters the mutagenicity of the intracellular environment. As discussed below, some genomic features can also have very localized effects on the mutation rate.

The reduction in fitness associated with a mutator allele is equal to the product of three terms: the excess genome-wide rate of production of deleterious mutations relative to the population mean (ΔU_D); the reduction in fitness per mutation (s); and the average number of generations that a mutation remains associated with the mutator ($\bar{t}$). The persistence time $\bar{t}$ is estimated by noting that an association between a mutator and an induced mutation is eliminated by selection at rate s per generation and by recombination at rate r (REF. 9).

Under asexual reproduction, the persistence time is the reciprocal of the mutational effect ($\bar{t} = 1/s$), and the selective disadvantage of a mutator allele (s_m) is simply equal to the increased rate of production of deleterious mutations ΔU_D. However, recombination weakens the selective disadvantage of a mutator allele by exporting mutations to other members of the population[9–11]. With free recombination ($r = 0.5$), mutant alleles are statistically uncoupled from their source in an average $\bar{t} = 2$ generations, so the mutator-allele disadvantage is $s_m \approx 2s\Delta U_D$, and no more than twice this with stronger linkage[6].