

## Opinion piece

## Why do species vary in their rate of molecular evolution?

**Despite hopes that the processes of molecular evolution would be simple, clock-like and essentially universal, variation in the rate of molecular evolution is manifest at all levels of biological organization. Furthermore, it has become clear that rate variation has a systematic component: rate of molecular evolution can vary consistently with species body size, population dynamics, lifestyle and location. This suggests that the rate of molecular evolution should be considered part of life-history variation between species, which must be taken into account when interpreting DNA sequence differences between lineages. Uncovering the causes and correlates of rate variation may allow the development of new biologically motivated models of molecular evolution that may improve bioinformatic and phylogenetic analyses.**

**Keywords:** molecular clock; phylogenetics; dating; mutation; substitution; population size

### 1. INTRODUCTION

Studying the correlates of variation in rate of molecular evolution is important for two reasons. First, rate variation is a window on molecular evolutionary processes: many theoretical models make predictions about rates that can be tested by analyses of rate variation within and between genomes. An examination of rates of molecular evolution is essential for asking fundamental questions about molecular evolution such as: are mutation rates shaped by selection? How closely is molecular evolution coupled to phenotypic evolution and speciation? Second, biologists are increasingly relying on DNA analysis in their research, yet most analytical methods make strong assumptions about rate variation between and within genomes, and results can be misleading if these assumptions are not met. Increasing our understanding of the causes and consequences of rate variation provides a platform for assessing the reliability of current methods and developing new methods that take the complexity of molecular evolution into account (Welch & Bromham 2005).

The genome contains a valuable source of information on evolutionary past and processes, as the divergence of lineages is incidentally recorded through the accumulation of heritable changes to DNA sequences. However, the way in which the genome evolves is affected by the kind of organism it is carried in. The purpose of this review is to consider the way that species' characteristics influence the rate of change

of DNA sequences. There are two basic phenomena that must be considered: the rate at which changes occur (mutation rate) and the rate at which these changes become incorporated into the species' shared genetic information (substitution rate).

### 2. MUTATION RATE VARIES BETWEEN SPECIES: DAMAGE AND COPY ERRORS

Nucleotide sequences can be permanently altered by physicochemical damage. Sometimes the damage itself changes the information, such as the deamination of cytosine (C) to uracil (U) so that it now pairs with A instead of G. In other cases, it is the repair of damaged bases that changes the sequence. For example, ultraviolet (UV) light can cause adjacent thymine residues to stick together. One way to repair this damage is by removing the damaged strand and replacing it with a new one, but the newly synthesized strand may contain incorrect bases (see Bromham 2008).

So the mutation rate due to damage is affected by two factors: the relative impact of mutagens; and the efficiency of damage repair. Both of these factors can vary between species. Some mutagens arise internally due to cellular processes such as metabolism. It has been suggested that species with higher metabolic rates generate more intracellular mutagens and thereby suffer a greater rate of DNA damage per unit time (e.g. Martin & Palumbi 1993), although the influence of such an effect on rates of molecular evolution has been disputed (see Lanfear *et al.* 2007; Galtier *et al.* 2009). Other mutagens come from the environment. For example, plants in areas of high environmental energy have faster rates of molecular evolution (Davies *et al.* 2004): one possible explanation is that high-energy environments have a direct mutagenic effect, perhaps through increased UV radiation (Wright *et al.* 2006).

The other source of mutations is DNA replication. Every time the genome is copied, there is a small chance of an error that changes the base sequence. So the mutation rate due to copy errors is determined by both the rate of error per copy and the number of copies made per unit time. Both of these factors may be influenced by species biology. For example, in many vertebrate lineages, species with shorter generation times have faster rates of molecular evolution, presumably because they copy their germline DNA more often per year (e.g. Bromham *et al.* 1996). The generation time effect may also apply to plants (e.g. Smith & Donoghue 2008), but has not been widely tested for other taxa.

The number of DNA replications per generation can also vary between species (Bartosch-Harlid *et al.* 2003). The numbers of cell generations taken to produce gametes can vary: for example, it takes fewer cell generations to make mouse ova than human ova. Number of DNA replications per generation can also vary with population structure and mating system. For example, eusocial bees and wasps have higher substitution rates than their non-social relatives, possibly because social queens produce vastly more eggs than non-social females, so copy their germline DNA thousands of times more per generation (Bromham & Leys 2005). Similarly, bird species with promiscuous mating systems have a more pronounced male mutation bias, suggesting that increased sperm

One contribution of 11 to a Special Feature on 'Whole organism perspectives on understanding molecular evolution'.

production results in a higher average number of male germline replications (Bartosch-Harlid *et al.* 2003). Because behavioural differences can influence the number of cell generations per organism generation, even closely related species can vary in the number of replication errors they accumulate, and thus can differ in their absolute mutation rates.

The per-replication mutation rate is affected by the efficiency of DNA repair, which determines how many copy errors persist uncorrected. Species can differ in the amount and efficiency of copy error detection and repair (Drake *et al.* 1998). Furthermore, efficiency of error correction can vary between individuals within a population due to differences in the genes that code for the repair machinery (e.g. Woodruff *et al.* 1984). This suggests DNA repair efficiency is open to selection. This is most obvious in bacterial populations where exposure to rapidly changing environments or strong selection pressures may select for 'mutators' that have lower DNA repair efficiency and therefore higher mutation rates (e.g. Denamur & Matic 2006). But selection may also shape rates of molecular evolution in eukaryotes (Baer *et al.* 2007).

### 3. SUBSTITUTION RATE VARIES BETWEEN SPECIES: SELECTION AND DRIFT

The number and type of mutations that go to fixation are determined by the balance between two forces, selection and drift, which can vary between species. The effectiveness of selection can vary across the genome, influenced by factors such as the recombination rate (Comeron *et al.* 2008), potentially leading to local increases in fixation of slightly deleterious mutations in regions of low recombination. But is it possible for selection to influence the genome-wide rates of substitution in a species, so that all (or nearly all) genes experience an elevated rate of change?

Some species may undergo a general relaxation of constraints due to a change in lifestyle. For example, parasites might not need to maintain functions provided by their host, such as nutrition, defence and dispersal. This could result in an increase in substitution rate across many genes, as many mutations that would be deleterious in a free-living organism become nearly neutral or neutral in the parasite. This may be why faster rates of molecular evolution have been reported in a number of parasitic taxa (e.g. Downton & Austin 1995; Duff & Nickrent 1997). Domestic breeds may provide another example of widespread relaxation of selection across the genome, because, similar to parasites, they have a 'host' that supplies nutrition, defensive structures, shelter and so on. (An alternative explanation is that selection on domestic or parasitic lineages favours novelty, so indirectly selects for increased generation of variation through mutation or recombination: e.g. Burt & Bell 1987; Denamur & Matic 2006; Dobney & Larson 2006.)

The power of selection to drive mutations to fixation is moderated by the sampling effects that occur every generation, as only a subset of alleles in the population are represented in subsequent generations. So one of the most pervasive effects on substitution rates is

population size (see Lynch 2007). In smaller populations, more substitutions go to fixation by drift, so if a substantial number of mutations are mildly deleterious, then species with small populations will have a higher overall substitution rate than those with larger populations. Effective population size can vary substantially between species. For example, species confined to islands have higher ratios of non-synonymous to synonymous substitutions than their mainland relatives, a result consistent with theoretical predictions of less effective selection in smaller populations (Woolfit & Bromham 2005). Lifestyle could also influence effective population size: for example endosymbiotic bacteria and fungi that are inherited without mixing maintain small populations and undergo frequent bottlenecks, thus have faster substitution rates than their free-living relatives (Woolfit & Bromham 2003). Since effective population size is highly labile, substitution rates can differ substantially between even closely related species (e.g. Petit & Barbadilla 2008).

### 4. THE GENOME AS A LIFE-HISTORY CHARACTER

All of these possible influences could blend together to create distinct differences between species in their rates of molecular evolution. For example, small-bodied mammal species tend to have faster rates of molecular evolution than their larger relatives (e.g. Martin & Palumbi 1993). Several potential causes of this pattern have been discussed, including both 'neutral' and adaptive processes that affect both the mutation and substitution rates. Smaller mammals have more generations per unit time, so should accumulate more DNA copy errors. But the relationship between generation time and substitution rate is not simple (Bromham *et al.* 1996). Mammal species differ in both the average number of cell generations per organism generation and the error rate per replication (due to both discrete differences in repair apparatus and continuous differences in repair efficiency: see Bromham *et al.* 1996). Furthermore, generation time co-varies with many other aspects of life history, so it may be that some other factor which scales with body size is the true causal factor. Small-bodied mammal species tend to live fast and die young, so if two related mammal species differ in body size, the smaller one is likely to have not only faster generations, but also a shorter lifespan, more offspring, higher metabolic rate and larger population size. All of these things could increase the rate of molecular evolution (Bromham *et al.* 1996; Nabholz *et al.* 2008; Welch *et al.* 2008).

Can rate of molecular evolution be considered a part of this life-history package? The costs of mutation may be higher for large, slow-maturing, long-lived mammal species. There is more opportunity for mutations to occur in a large organism which has more cells (and therefore more genome copies), takes more cell generations to produce gametes, and must maintain its body through a prolonged period of immaturity and long reproductive lifespan. And each mutation may have a higher cost. Because large mammals generally have fewer offspring, a single mutation in the germline could

destroy a greater percentage of its offspring. Plus, there are more somatic cells which could undergo disabling somatic mutations that reduce reproductive output of that individual. Furthermore, since large mammals maintain smaller populations, more slightly deleterious mutations will go to fixation by drift despite their negative effect on fitness. So large-bodied mammal species may have more opportunity for mutations to occur, each mutation has a potentially greater cost to fitness, and more deleterious mutations will become substitutions. Therefore, there should be stronger selection pressure to reduce mutation rates in larger species, so they may invest more in DNA copy fidelity and repair mechanisms, despite the metabolic and time costs involved. Like any other life-history trait, the rate of molecular evolution represents a balance between competing needs that vary according to a species size, reproductive strategies and lifestyle.

## 5. CONCLUSIONS

We need to recognize that the genome is the central adaptation of any organism. As with any adaptation, it is shaped by trade-offs between competing needs and processes. This does not limit the usefulness of DNA sequences as a source of information about evolution, but it does complicate extracting that information. While analytical methods are increasingly taking lineage-specific rate variation into account, many assume that rates are drawn randomly from a distribution, or evolve by a purely stochastic process. But rates can evolve in concert with species characteristics, which may create complex patterns of changing rates along phylogenies.

The growing body of available DNA sequences, life-history traits and ecological information for a large number of species now makes it possible to systematically explore the causes and correlates of rate variation. In addition to enlightening us on the processes of molecular evolution, information on the way species traits and molecular evolution interact may assist the construction of phylogeny or estimation of molecular dates if we can develop predictive models of rate change based on these phenotypic or ecological factors.

I am indebted to Rob Lanfear, John Welch and Meg Woolfit for their altruism.

Lindell Bromham\*

*Centre for Macroevolution and Macroecology, Botany and Zoology, School of Biology, Australian National University, Canberra, Australian Capital Territory 0200, Australia*

\*lindell.bromham@anu.edu.au

- Baer, C., Miyamoto, M. M. & Denver, D. R. 2007 Mutation rate variation in multicellular eukaryotes: causes and consequences. *Nat. Rev. Genet.* **8**, 619–631. (doi:10.1038/nrg2158)
- Bartosch-Harlid, A., Berlin, S., Smith, N. G. C., Iler, A. P. & Ellegren, H. 2003 Life history and the male mutation bias. *Evolution* **57**, 2398–2406. (doi:10.1554/03-036)
- Bromham, L. 2008 *Reading the story in DNA: a beginner's guide to molecular evolution*. Oxford, UK: Oxford University Press.
- Bromham, L. & Leys, R. 2005 Sociality and rate of molecular evolution. *Mol. Biol. Evol.* **22**, 1393–1402. (doi:10.1093/molbev/msi133)
- Bromham, L., Rambaut, A. & Harvey, P. H. 1996 Determinants of rate variation in mammalian DNA sequence evolution. *J. Mol. Evol.* **43**, 610–621. (doi:10.1007/BF02202109)
- Burt, A. & Bell, G. 1987 Mammalian chiasma frequencies as a test of two theories of recombination. *Nature* **326**, 803–805. (doi:10.1038/326803a0)
- Comeron, J. M., Williford, A. & Kliman, R. M. 2008 The Hill–Robertson effect: evolutionary consequences of weak selection and linkage. *Heredity* **100**, 19–31. (doi:10.1038/sj.hdy.6801059)
- Davies, T. J., Savolainen, V., Chase, M. W., Moat, J. & Barraclough, T. G. 2004 Environmental energy and evolutionary rates in flowering plants. *Proc. R. Soc. Lond. B* **271**, 2195–2200. (doi:10.1098/rspb.2004.2849)
- Denamur, E. & Matic, I. 2006 Evolution of mutation rates in bacteria. *Mol. Microbiol.* **60**, 820–827. (doi:10.1111/j.1365-2958.2006.05150.x)
- Dobney, K. & Larson, G. 2006 Genetics and animal domestication: new windows on an elusive process. *J. Zool.* **269**, 261–271. (doi:10.1111/j.1469-7998.2006.00042.x)
- Dowton, M. & Austin, A. D. 1995 Increased genetic diversity in mitochondrial genes is correlated with the evolution of parasitism in the Hymenoptera. *J. Mol. Evol.* **41**, 958–965. (doi:10.1007/BF00173176)
- Drake, J., Charlesworth, B., Charlesworth, D. & Crow, J. 1998 Rates of spontaneous mutation. *Genetics* **148**, 1667–1686.
- Duff, R. J. & Nickrent, D. L. 1997 Characterization of mitochondrial small-subunit ribosomal RNAs from holoparasitic plants. *J. Mol. Evol.* **45**, 631. (doi:10.1007/PL00006267)
- Galtier, N., Jobson, R. W., Nabholz, B., Glémin, S. & Blier, P. U. 2009 Mitochondrial whims: metabolic rate, longevity and the rate of molecular evolution. *Biol. Lett.* **5**, 413–416. (doi:10.1098/rsbl.2008.0662)
- Lanfear, R., Thomas, J. A., Welch, J. J., Brey, T. & Bromham, L. 2007 Metabolic rate does not calibrate the molecular clock. *Proc. Natl Acad. Sci. USA* **104**, 15 388–15 393. (doi:10.1073/pnas.0703359104)
- Lynch, M. 2007 *The origins of genome architecture*. Sunderland, MA: Sinauer Associates.
- Martin, A. P. & Palumbi, S. R. 1993 Body size, metabolic rate, generation time and the molecular clock. *Proc. Natl Acad. Sci. USA* **90**, 4087–4091. (doi:10.1073/pnas.90.9.4087)
- Nabholz, B., Glémin, S. & Galtier, N. 2008 Strong variations of mitochondrial mutation rate across mammals: the longevity hypothesis. *Mol. Biol. Evol.* **25**, 120–130. (doi:10.1093/molbev/msm248)
- Petit, N. & Barbadilla, A. 2008 Selection efficiency and effective population size in *Drosophila* species. *J. Evol. Biol.* **22**, 515–526. (doi:10.1111/j.1420-9101.2008.01672.x)
- Smith, S. A. & Donoghue, M. J. 2008 Rates of molecular evolution are linked to life history in flowering plants. *Science* **322**, 86–89. (doi:10.1126/science.1163197)
- Welch, J. J. & Bromham, L. 2005 Estimating molecular dates when rates vary. *Trends Ecol. Evol.* **20**, 320–332. (doi:10.1016/j.tree.2005.02.007)
- Welch, J. J., Bininda-Emonds, O. R. P. & Bromham, L. 2008 Correlates of substitution rate variation in mammalian protein-coding sequences. *BMC Evol. Biol.* **8**, 53. (doi:10.1186/1471-2148-8-53)
- Woodruff, R. C., Thompson, J. J. J., Seeger, M. A. & Spivey, W. E. 1984 Variation in spontaneous mutation

- and repair in natural population lines of *Drosophila melanogaster*. *Heredity* **53**, 223–234. (doi:10.1038/hdy.1984.78)
- Woolfit, M. & Bromham, L. 2003 Increased rates of sequence evolution in endosymbiotic bacteria and fungi with small effective population sizes. *Mol. Biol. Evol.* **20**, 1545–1555. (doi:10.1093/molbev/msg167)
- Woolfit, M. & Bromham, L. 2005 Population size and molecular evolution on islands. *Proc. R. Soc. B* **272**, 2277–2282. (doi:10.1098/rspb.2005.3217)
- Wright, S. D., Keeling, J. & Gillman, L. 2006 The road from Santa Rosalia: a faster tempo of evolution in tropical climates. *Proc. Natl Acad. Sci. USA* **103**, 7718–7722. (doi:10.1073/pnas.0510383103)