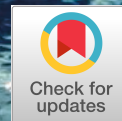


INSIGHTS

PERSPECTIVES



EVOLUTIONARY BIOLOGY

The raw material of evolution

Estimates of whale mutation rates contribute to understanding evolutionary processes

By **A. Rus Hoelzel¹** and **Michael Lynch²**

Evolution happens when the code of life, DNA, is changed by the process of mutation. Mutations include deletions and insertions, rearrangements, and transpositions (moving DNA). However, it is the rate of point mutations, which affect a single site in the chain of nucleotides that make up DNA, that is most often considered. Knowing the rate and pattern of mutation is essential to understanding the process of evolution. This knowledge increases understanding of natural selection and has applications such as tracking demography and dating phylogenies, but calculating the mutation rate is not straightforward. On page

990 of this issue, Suárez-Menéndez *et al.* (1) use parent-offspring trios and genome sequences to estimate rates of point-mutational change from one generation to the next in four species of baleen whales in the North Atlantic: humpback (*Megaptera novaeangliae*), blue (*Balaenoptera musculus*), fin (*Balaenoptera physalus*), and bowhead (*Balaena mysticetus*).

Generating sufficient data from such inaccessible animals is a challenge. However, Suárez-Menéndez *et al.* sampled extensively enough to identify parent-offspring trios—five trios for the humpback whale and one for each of the other three species—using a relatively inexpensive and rapid genetic method for kinship screening. Their results matter because a mutation rate cannot simply be calculated for one species, for which this can be easily done, and then applied more generally to other taxa. There is no universal mutation rate. There is not even a single mutation

rate for a given taxon—it varies across individuals and among nuclear, mitochondrial, and chloroplast genomes. The mutation rate is subject to natural selection like all traits (2, 3), and mutation rates in nuclear genomes vary 10,000-fold across the tree of life and 40-fold among vertebrates (3).

Estimating mutation rates in very large mammals is of interest in part to help address how species with long generation times (average time between two consecutive generations) avoid high incidence of somatic diseases such as cancer. Suárez-Menéndez *et al.* estimated the nuclear point-mutation rate for the four species of baleen whales combined to be 1.11×10^{-8} , which is high compared with earlier estimates that were based on phylogenies. Therefore, in cetaceans a lower than expected incidence of cancer may be due to selection, as suggested by Tejada-Martínez *et al.* (4) and others, rather than a slower mutation rate.

After a mutation occurs, the individual

¹Department of Biosciences, Durham University, Durham, UK. ²Biodesign Center for Mechanisms of Evolution, Arizona State University, Tempe, AZ, USA.
Email: a.r.hoelzel@durham.ac.uk



Humpback whales (*Megaptera novaeangliae*) are often inaccessible, which makes assessment of their mutation rate challenging.

be anticipated and modeled (6), but there are uncertainties involved. Determining the mutation rate with trio analyses avoids these issues, because time is too short for the loss of anything other than a dominant lethal mutation.

Knowledge of the mutation rate can reveal other important aspects of the biology of a species, for example, demography. The genetic effective population size (N_e) is an idealized approximation of the average number of individuals that produce successful progeny per generation, which governs the level of noise in evolutionary processes. Suárez-Menéndez *et al.* used linkage disequilibrium analysis (nonrandom association of loci) and calculated the N_e of humpback whales before commercial whaling began to be 5,700, though, as the authors point out, there are caveats to these estimations. N_e is generally smaller than the total number of reproductive adults in the population, owing to variance in reproductive success among individuals, demographic fluctuations in population size, and deleterious mutations (which can purge variation from linked chromosomal regions). The ratio of N_e to the census population size varies widely (7) and, for the humpback whale, may be roughly 0.1 (8).

All methods for estimating demographic trajectories depend critically on accurate measures of both the mutation rate and the generation time. An alternative method for calculating this trajectory is based on coalescence (the time at which evolutionary lineages come together) and the knowledge that the interval between these coalescence points varies with N_e [(9), compare with (10)]. For example, the skyline model can estimate these demographic trajectories and incorporate ancient DNA to estimate the relevant mutation rate [but see (11, 12)]. These rates may be somewhat slower than trio-based rates (6) but can provide remarkably close correlation with known historical events (13) and are typically faster than substitution rates calculated from fossil-calibrated phylogenies (1, 5).

A long-term average N_e can be calculated from the diversity of the population at putatively neutral sites within the genome (the behavior of which is not biased by selection) and the mutation rate. A notable negative correlation between taxon-specific nuclear mutation-rate estimates and their long-term N_e is found when organisms are compared across the entire tree of life (14), and now there are enough data to see this trend emerging among mam-

mals. Estimates of mammalian mutation rates (per generation) range approximately fourfold, from 0.4×10^{-8} (pig) to 1.6×10^{-8} (gray mouse lemur), with cetaceans having relatively high estimated mutation rates, in the range of 0.9×10^{-8} to 1.4×10^{-8} (1, 3). The observed negative scaling is consistent with the drift-barrier hypothesis (15), which postulates that as N_e declines, drift reduces the ability of selection to purge mutator alleles. When N_e is large, selection is more efficient at reducing the mutation rate. Other factors that may influence vertebrate mutation rates include parental age, species-level fecundity, a higher mutation rate in males than females [in birds and mammals (3)], and genome size (14).

Several uncertainties remain with respect to the new data from Suárez-Menéndez *et al.* as well as recent estimates from other species of vertebrates (3). Estimates of N_e that are derived by dividing standing levels of variation at putatively neutral sites by mutation rate are subject to biases, with the bias direction depending on the relative magnitude of sampling error that is associated with estimates of nucleotide variation and mutation rates. Additionally, most mammalian mutation-rate estimates are based on very small numbers of trios (often only one), although it is known that mutation rates can vary by a factor of at least two among individuals within a species (14). Thus, new estimates such as the one presented by Suárez-Menéndez *et al.*, which have been notably lacking for long-lived species other than humans, represent progress toward a fuller understanding of the rate of origin of the raw material of evolution. At the same time, caution is needed when dating past evolutionary events with present-day mutation rates and drawing inferences about the biological basis for taxon-specific mutation-rate differences. ■

REFERENCES AND NOTES

1. M. Suárez-Menéndez *et al.*, *Science* **381**, 990 (2023).
2. W. Wei *et al.*, *Nat. Commun.* **13**, 4752 (2022).
3. L. A. Bergeron *et al.*, *Nature* **615**, 285 (2023).
4. D. Tejada-Martínez *et al.*, *Proc. R. Soc. Lond. Ser. B* **288**, 20202592 (2021).
5. S.-Y. W. Ho *et al.*, *Syst. Biol.* **60**, 366 (2011).
6. D. Graur, *Molecular and Genome Evolution* (Sinauer, 2016).
7. S. Hoban *et al.*, *Biol. Conserv.* **248**, 108654 (2020).
8. A. L. Cypriano-Souza, T. F. da Silva, M. H. Engel, S. L. Bonatto, *Genet. Mol. Biol.* **41**, 253 (2018).
9. S.-Y. W. Ho, B. Shapiro, *Mol. Ecol. Resour.* **11**, 423 (2011).
10. K. V. Parag, O. G. Pybus, C.-H. Wu, *Syst. Biol.* **71**, 121 (2021).
11. P. Johnri *et al.*, *PLOS Biol.* **20**, e3001669 (2022).
12. S. Möller, L. du Plessis, T. Stadler, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 4200 (2018).
13. M. de Bruyn, A. R. Hoelzel, G. R. Carvalho, M. Hofreiter, *Trends Ecol. Evol.* **26**, 405 (2011).
14. M. Lynch *et al.*, *Nat. Rev. Genet.* **17**, 704 (2016).
15. M. Lynch, *Genome Biol. Evol.* **3**, 1107 (2011).

10.1126/science.adk0121

with the mutation may fail to survive or reproduce because of impairment due to the mutation (purifying selection) or simply by failing to reproduce by chance (genetic drift). Therefore, a comparison of populations or species after 100 or 1000 generations will reflect net genetic differentiation that results from de novo variation introduced by mutation, as estimated by trio analysis, which is then modified by subsequent forces of selection and drift. Eventually, the measured rate will reflect new variants shared by all members of the population (the substitution rate), which can be more than an order of magnitude lower than the de novo mutation rate, but there may be intermediate rates for a period of time (5). A further factor that is frequently encountered in the relatively rapidly changing mitochondrial DNA genomes of mammals is saturation. This occurs when mutation changes one nucleotide to another, and then back again. This can