Mutation Rates: Simpler than we thought?

John F.Y. Brookfield

School of Life Sciences

University of Nottingham

University Park

Nottingham NG7 2RD

United Kingdom

Summary

Mutation rate variation is often explained by varying optimal rates, or through effective population sizes determining the effectiveness of selection. But a rate difference between humans and owl monkeys is now explained mechanistically as a consequence of differing reproductive longevities.

Mutation creates the variation that underlies all evolutionary change. While mutations are far more likely to be harmful than beneficial, evolution's requirement for mutational input leads to the conclusion that the optimal rate of mutation cannot be zero. Due to mutation's association with evolution, one can easily imagine that mutational rates may themselves be evolutionarily optimal, and adjusted to the nuances of each species' requirements. However, a new paper from Thomas *et al*. [1] suggests that an observed difference in mutation rates between humans and a species of owl monkey (*Aotus nancymaae*), with the monkey's rate being just 68% of human's, is, in fact, easily explained on mechanistic grounds as a function of the reproductive longevity of this species.

The measurement of mutation rates at the DNA level is difficult. We can measure the rate of mutation when mutation creates a selectable phenotype in a microbial system, but it is hard to extrapolate to the rest of the DNA sequences in the genome, since mutation rates can differ

between loci. A more indirect route is to infer the mutation rate from the rate of evolution. For genomic sites where variation is neutral, the evolutionary rate per generation will be the same as the mutation rate per generation, since neutral mutational changes are sampled in an unbiased way by the evolutionary process. Thus, using interspecies comparisons to look at the rate of DNA evolution per year at, for example, fourfold degenerate codon positions, estimates of the mutation rate per generation can be produced, provided we trust our estimates of the times of evolutionary divergences and of generation times, and have confidence that our assumption of complete neutrality is correct. But, in this decade, it has become possible, using whole genome sequencing, to look at reproductive trios, and to identify, directly, mutations as variants in offspring that are absent in both parents [2,3].

If mutation rates are optimal, what determines what that optimum will be? It is the possibility of adaptive mutations that raises the optimal mutation rate above zero. As the rate of mutation to selectively advantageous alleles per gamete is μ_s, the rate of selectively advantageous mutation at the population level for a diploid is $2N_e\mu_s$ (where N_e is the effective population size). The probability of selective spread of an advantageous allele is approximately $2s$ [4] (when s, the selective advantage, is small), giving a diploid rate of adaptive evolution of $4N_e s\mu_s$. So a high μ_s gives a high rate of adaptive evolution, and the advantage of mutation can thereby be quantified. But μ_s is not a biochemical constant. Rather, a species in a constant environment might have found the mutations needed to adapt to this environment, and μ_s will be zero or very small. Alternatively, in a recently changed environment, μ_s will be larger, not because the overall mutation rate has increased, but because, in a population poorly adapted to a changed environment, the proportion of mutations that are selectively advantageous will be enhanced.

What are the consequences for the evolution of mutation rates? The higher μ_s is, and the higher the selective advantage associated with favourable mutations, the stronger will be the selection in

favour of a mutator allele [5]. Indeed, there is some evidence that under stressful conditions (which might be associated with poor adaptation to the environment), rates of mutation are enhanced in *Escherichia coli* [6]. Also, there is some evidence that mutation rates are lower in genes subject to the strongest purifying selection [7]. But selection for a trans-acting mutator allele requires it to be found in higher fitness individuals than its alternative allele, and this association is broken down by recombination. In the long-term evolution experiment [8, 9] with *E. coli*, of twelve populations, six evolved to a mutator phenotype, as this was favoured by the lack of recombination and also from the higher μ_s and mean *s*, expected in populations adapting to a new environment.

An alternative view of mutation rate is that, in typical environments, the supply of mutations does not limit adaptive evolution, and selection is typically for anti-mutator alleles [10], as a result of such anti-mutator alleles being genomically associated with fewer new harmful alleles. But this selection will be weak. Whether it is effective will depend on the effective population size. If the product of the selection in favour of an anti-mutator allele and the effective population size is less than one, whether or not an anti-mutator allele will fix will be dependent upon genetic drift. In this way, there will be an effective drift barrier preventing evolution of very low per generation mutation rates when the effective population size is low.

The owl monkey study sequenced genomes from 30 individuals from six multi-genome pedigrees [1], and found the overall per generation mutation rate to be reduced by 32% relative to humans. There is a clear effect of paternal age on mutation rates, but the data suggest that it is the number of years experienced post-puberty in the males that is linearly related to mutation rate, rather than total age. And if human mutation rates, as a function of years post-puberty, are used to predict the owl monkey's overall mean mutation rate, allowing for the different life history of the monkey, the agreement with the monkey data is very close.

Thus, here, it appears that a difference in mutation rates between species, for which an adaptive explanation might potentially have been sought, reflects a difference in life history between species, with the lowered rate being a consequence simply of the lowered number of years post-puberty during which germline mutations have accumulated in the monkey. The authors approach can also be shown to yield accurate mutation rate predictions for the chimpanzee. But it is not clear whether this result can be extended outside the primates. (As a new world monkey, the owl monkey is phylogenetically distant from the apes.) This thus represents a new null hypothesis with which observed mutation rates (and there will surely be forthcoming studies of many species using whole genome sequencing approaches) can be interpreted. We no longer expect equal rates, but rates predictable from the number of years post-puberty experienced by the parents. Thus interesting exceptions, where adaptive explanations could be contemplated, would be those which failed to agree with this null. Another consequence of the model will be that (provided we have enough life history data for the relevant species) the "generation time effect", where short-lived species tend to have higher rates of molecular evolution, can be re-examined in a more detailed and predictive way.

References:

1. Thomas, G.W.C., Wang, R.J., Puri, A., Harris, R.A., Raveendran, M., Hughes, D., Murali, S. Williams, L., Doddapaneni, H., Muzny, D., Gibbs, R., Abee, C., Galinski, M.R., Worley, K.C., Rogers, J., Radivojac, P., and Hahn, M.W. (2018) Reproductive longevity predicts mutation rates in primates. Curr Biol. ?, ?

2. Cartwright, R.A., Hussin, J., Keebler, J.E., Stone, E.A., and Awadalla, P. (2012). A family-based probabilistic method for capturing de novo mutations from high-throughput short-read sequencing data. Stat. Appl. Genet. Mol. Biol. 11, 11.

3. Kong, A., Frigge, M.L., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., Gudjonsson, S.A., Sigurdsson, A., Jonasdottir, A., Jonasdottir, A., Wong, W.S., Sigurdsson, G., Walters, G.B., Steinberg, S., Helgason, H., Thorleifsson, G., Gudbjartsson, D.F., Helgason, A., Magnusson, O.T.,

Thorsteinsdottir, U., and Stefansson, K., (2012). Rate of de novo mutations and the importance of father's age to disease risk. Nature 488, 471–475.

4. Haldane J.B.S. (1927) The mathematical theory of natural and artificial selection. Proc. Camb. Philos. Soc. 23, 838–844.

5. Taddei, F., Radman, M., Maynard-Smith, J., Toupance, B., Gouyon, P.H., and Godelle, B., (1997) Role of mutator alleles in adaptive evolution. Nature 387, 700-702.

6. Bjedov, I., Tenaillon, O., Gérard, B., Souza, V., Denamur, E., Radman, M., Taddai, F., and Matic, I., (2003) Stress-induced mutagenesis in bacteria. Science 300, 1404-1409.

7. Martincorena, I., Neshasayee, A.S.N., Luscombe, N.M., (2012) Evidence of non-random mutation rates suggests an evolutionary risk management strategy. Nature 485, 95-98.

8. Couce, A., Caudwell, L.V., Feinauer, C., Hindré, T., Feugeas, J-P., Weigt, M., Lenski, R.E., Schneider, D., and Tenaillon, O. (2017) Mutator genomes decay, despite sustained fitness gains, in a long-term experiment with bacteria. Proc Natl. Acad Sci USA 114, E9026-E9035.

9. Wielgoss, S., Barrick, J.E., Tenaillon, O., Wiser, M.J., Dittmar, W.J., Cruveiller, S., Chane-Woon-Ming, B., Médigue, C, Lenski, R.E., and Schneider, D., (2013) Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load. Proc Natl Acad Sci USA 110, 222–227.

10. Lynch, M., Ackerman, M.S., Gout, J.-F., Long, H., Sung, W., Thomas, W.K., and Foster, P.L., (2016) Genetic drift, selection, and the evolution of the mutation rate. Nature Reviews Genetics 17, 704-714.

Figure 1: A stylised illustration of the model of mutation rates from Thomas et al. [1]. The three lines represent the increase in mutation rates with age in three theoretical organisms. The blue line represents a species with puberty at four years, the red line a species with puberty at seven years, while the green line is a species with puberty at twelve years. The mutation rate at puberty is the same in all species, as is the slope of the relationship between mutation rate and the number of years post-puberty. But, in this model, it is possible for a species with a lower mean age at reproduction (shown by the blue square) to have a higher mutation rate than a species with a higher mean age at reproduction (the green square).