DOI: 10.1534/genetics.108.098905

Letter to the Editors

Waiting Longer for Two Mutations

Michael J. Behe[1]

Department of Biological Sciences, Lehigh University, Bethlehem, Pennsylvania 18105

Manuscript received November 19, 2008

IN the Abstract of their recent article, "Waiting for Two Mutations: With Applications to Regulatory Sequence Evolution and the Limits of Darwinian Evolution" (GENETICS **180:** 1501–1509, 2008), Durrett and Schmidt write that one of their aims is "to expose flaws in some of Michael Behe's arguments concerning mathematical limits to Darwinian evolution." Their effort, however, is itself seriously flawed.

They develop a population genetics model to estimate the waiting time for the occurrence of two mutations, one of which is premised to damage an existing transcription-factor-binding site, and the other of which creates a second, new binding site within the nearby region from a sequence that is already a near match with a binding site sequence (for example, 9 of 10 nucleotides already match). They model two separate cases: where the mutation of the initial transcription site is neutral and where it is deleterious. Toward the end of the article they criticize my observation of the rarity of two amino acid substitutions in the chloroquine-resistant form of the protein PfCRT from the viewpoint of their model. Citing malaria literature sources (WHITE 2004) I had noted that the *de novo* appearance of chloroquine resistance in *Plasmodium falciparum* was an event of probability of 1 in 10^{20}. I then wrote that "for humans to achieve a mutation like this by chance, we would have to wait 100 million times 10 million years" (BEHE 2007) (because that is the extrapolated time that it would take to produce 10^{20} humans). DURRETT and SCHMIDT (2008, p. 1507) retort that my number "is 5 million times larger than the calculation we have just given" using their model (which nonetheless gives a prohibitively long waiting time of 216 million years).

Their criticism compares apples to oranges. My figure of 10^{20} is an empirical statistic from the literature; it is not, as their calculation is, a theoretical estimate from a population genetics model. Generally, when the results of a simple model disagree with observational data, it is an indication that the model is inadequate. Furthermore, DURRETT and SCHMIDT (2008) err in several ways in applying their model to the PfCRT data:

1. For the rate of the first mutation, Durrett and Schmidt use a value estimated for the alteration of a transcription-factor-binding site, where any of 10 nucleotides could be changed. In the case of the protein, however, it is likely that a particular nucleotide of a particular amino acid residue's codon must be changed. This introduces a 30-fold underestimate of the waiting time.
2. They use the model that they developed for an initial neutral mutation, but it is likely that the initial protein point mutation is deleterious. If it is strongly deleterious, their calculation could be low by many orders of magnitude, as their own model for deleterious mutations shows.

Finally, their model is incomplete on its own terms because it does not take into account the probability of one of the nine matching nucleotides in the region that is envisioned to become the new transcription-factor-binding site mutating to an incorrect nucleotide before the 10th mismatched nucleotide mutates to the correct one. Since the mutation rates for all nucleotides are presumably of the same order, this introduces an independent underestimate of a factor of nine for their own model. In applying the model to the PfCRT protein, this overlooked factor is much more severe. If after the first "correct" mutation, a mutation occurs in an amino acid codon other than the needed second one, there is a strong chance that it would damage the activity of the protein. Since the gene for the protein is >1000 nucleotides in length, this introduces an underestimate of several orders of magnitude.

When DURRETT and SCHMIDT (2008) noted that their own estimate of obtaining two needed mutations in humans was an unrealistically long 216 million years, they pondered that, if the regulatory "neighborhood" for a gene was not just 1 kb but 1000 kb, then "we expect to find 16 copies" of the already complete transcription-factor-binding site there. But if many useful binding sites already exist in the neighborhood, why model the appearance of yet another one? The difficulty with

[1]*Address for correspondence:* Department of Biological Sciences, 111 Research Dr., Lehigh University, Bethlehem, PA 18105.
E-mail: mjbl@lehigh.edu

models such as Durrett and Schmidt's is that their biological relevance is often uncertain, and unknown factors that are quite important to cellular evolution may be unintentionally left out of the model. That is why experimental or observational data on the evolution of microbes such as *P. falciparum* are invaluable—because they can constrain our models. Whatever we speculate about what may be the usefulness of a new transcription-factor-binding site, gene duplication, meiotic recombination, protein domain swap, or anything else, none of them were of much use in helping the malarial parasite fend off an evolutionary challenge. The data show that in 10^{20} chances only several point mutations in PfCRT were useful to it in effectively combating chloroquine.

LITERATURE CITED

Durrett, R., and D. Schmidt, 2008 Waiting for two mutations: with applications to regulatory sequence evolution and the limits of darwinian evolution. Genetics **180:** 1501–1509.

White, N. J., 2004 Antimalarial drug resistance. J. Clin. Invest. **113:** 1084–1092.

Behe, M. J., 2007 *The Edge of Evolution: The Search for the Limits of Darwinism.* Free Press, New York.