Genetics: Published Articles Ahead of Print, published on February 2, 2009 as 10.1534/genetics.109.100800

Reply to Michael Behe

We showed that the mean waiting time for two mutations to occur in the same individual, one with probability u_1 and another with probability u_2 (when the first mutation is neutral), is $1/2Nu_1 u_2^{1/2}$. The square root on the second factor is an important insight from our calculation and is the main difference between our theory and Behe's naïve calculations, which assume that the two mutations must occur almost simultaneously. Our results show that there are of order $1/u_2^{1/2}$ individuals with the first mutation before the second one occurs (see the sketch of the proof of Theorem 1 on page 1503).

"… 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 tenth, mismatched codon mutates to the correct one." This conclusion is simply wrong since it assumes that there is only one individual in the population with the first mutation. There are of order $1/u_2^{1/2}$ individuals with the first mutation before the second one occurs, and since this event only removes one individual from the group with the first mutation it has no effect on the waiting time.

Behe is not alone in making this type of mistake. When Evelyn Adams won the New Jersey lottery on October 23, 1985 and again on February 13, 1986, newspapers quoted odds of 17.1 trillion to 1. That assumes the winning person and the two lottery dates are specified in advance, but at any point in time there is a population of individuals who have won the lottery and have a chance to win again, and there are many possible pairs of dates on which this event can happen. The probability it happens in one lottery during one year is about 1 in 200 (see Example 2.33 in Durrett (2009) for details).

"For the rate of the first mutation they use a value estimated for the alteration of a transcription factor binding site, where any of ten 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 thirty-fold underestimate of the waiting time." Behe is right on this point. This divides our previously computed overestimate of 5 million by 30.

"They use the model 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." If the first mutation is mildly deleterious (a fitness loss of order $u_2^{1/2}$), then the waiting time is increased by a factor of 2 or 3. If the loss of fitness were 0.1, then the mean waiting time would be $1/(20Nu_1u_2)$. We leave it to biologists to debate whether the first PfCRT mutation is that strongly deleterious.

"My figure of 10^{20} [the odds of a malaria parasite developing resistance to chloroquine] *is an empirical statistic from the literature; it is not, like their calculation, a theoretical estimate from a population genetics model."* We disagree that Behe's result is an empirical fact. It is clearly impossible to know the number of times the double mutation has occurred, so in order to infer that from the number of times the mutation has avoided extinction in an individual and risen to a frequency where it can be noticed in a subpopulation requires a model, which we have provided.

Finally, Behe notes that for one prespecified pair of mutations in one gene in humans with the first one neutral we get a "prohibitively long waiting time" of 216 million years. However, there are at least 20,000 genes in the human genome and for each gene tens if not hundreds of pairs of mutations that can occur in each one. Our results show that the waiting time for one pair of mutations is well

approximated by an exponential distribution. If there are k nonoverlapping possibilities for double mutations, then by an elementary result in probability, the waiting time to the first occurrence is the minimum of k independent exponentials and hence has an exponential distribution with a mean that is divided by k. From this we see that in the case in which the first mutant is neutral or mildy deleterious double mutations can easily have caused a large number of changes in the human genome since our divergence from chimpanzees. Of course if the first mutant already confers an advantage then such changes are easier.

R. Durrett (2009) Elementary Probability for Applications. To be published by Cambridge U. Press. Latest draft available at www.math.cornell.edu/~durrett/