HHMI
HOWARD HUGHES MEDICAL INSTITUTE

ASK A SCIENTIST

Q&A

Go BACK

QUESTION

Submitted by

Anonymous, Virginia, USA

I have just watched the Holiday Lectures on evolution DVD. I was especially interested in those portions of the DVD related to the increase in the proportion of mice with black fur relative to the proportion of mice with sandy fur once a mutation has resulted in the birth of a mouse with black fur. I have made an attempt to create a model for the changing proportions that would give the results shown in the DVD.

To do this I made the following assumptions:

- There is a carrying capacity of 5,000 for the entire mouse population.
- The population of mice with dark fur would be modeled nicely by a logistic function based on a carrying capacity of 5,000 and a per capita growth rate of 0.01% per generation.
- I used the dark fur mouse population model to create a proportion model by simply dividing the increasing dark fur population numbers by 5,000.

The graph of my dark fur proportion function had the classic logistic shape of the graph in the DVD. However, according to my model, the proportion did not reach the 95% level until about generation 1,146 instead of generation 1,003 according to the model used to generate the graph on the DVD.

How should I alter my assumptions to more closely re-create the model shown in the DVD lecture?

Science **PROFILE**

CHRIS HITTINGER, Ph.D.
Maclyn McCarty Fellow of the Helen Hay Whitney Foundation
University of Colorado Health Sciences Center

ANSWER

Provided by

Chris Hittinger
Maclyn McCarty Fellow of
the Helen Hay Whitney Foundation,
Department of Biochemistry and Molecular Genetics,
University of Colorado Health Sciences Center
(former HHMI predoctoral fellow)

The equation for the generation when a particular allele frequency is reached for a dominant allele experiencing positive selection is:

$$t = [\ln (p_t/q_t) - \ln (p_0/q_0) + 1/q_t - 1/q_0]/s$$

where t = the number of generations to reach the desired frequency; p_t = desired black allele frequency; $q_t = 1 - p_t$ = desired tan allele frequency; p_0 = starting black allele frequency; $q_0 = 1 - q_0$ = starting tan allele frequency; s = selection coefficient (favoring black mice); and ln is the natural log function.

Although the full derivation is somewhat laborious, this is a standard population genetics equation manipulated algebraically from the textbook cited below. Of course, it can be further manipulated algebraically to solve for any of the components, instead of t.

We ignore the stochastic effects of genetic drift (and hence population size) and assume random mating. The black allele is completely dominant to the tan allele. Sean Carroll's presentation (in the DVD) shows an invading population of 100 mice with a single (heterozygous) black mouse (1 black allele and 1 tan allele) and 99 tan mice (198 tan alleles).

Therefore, p_0 = 1 black allele/200 total alleles = 0.005.

$q_0 = 1 - p_0 = 1 - 0.005 = 0.995$.

We are looking for the point at which 95% of the mice are black, which means 5% of mice will be tan. Since tan mice must be homozygous, $q_t^2 = 0.05$, which means $q_t = (0.05)^{1/2} = 0.2236$.

Since $q_t = 1 - p_t$, it is also true that $p_t = 1 - q_t = 1 - 0.2236 = 0.7764$, meaning 95% of mice will be black when the black allele frequency has reached this proportion.

Plugging in all of our values with a selection coefficient of 0.01 (as noted in the lecture, the real one estimated in the wild is much higher) gives us:

$t = [\ln (p_t/q_t) - \ln (p_0/q_0) + 1/q_t - 1/q_0]/s =$

$[\ln (0.7764/0.2236) - \ln (0.005/0.995) + 1/0.2236 - 1/0.995]/0.01 =$

$[\ln (3.4723) - \ln (0.005025) + 4.4723 - 1.0050]/0.01 =$

$[1.2448 - (-5.2933) + 4.4723 - 1.0050]/0.01 = 10.005/0.01 = 1000.5$ generations, which is pretty close to the 1000 generations modeled in the DVD.

Reference

Hartl, D.L., and Clark, A.G. *Principles of Population Genetics*, 3rd ed. Sunderland, MA: Sinauer Associates, 1997.

02/03/09 21:38