MCB137L/237L: Physical Biology of the Cell
Spring 2022
Homework 3
(Due 2/10/22 at 11:00am)

Hernan G. Garcia

"Their exercises are unbloody battles, and their battles bloody exercises." - Flavius Josephus on Roman legionares

1a. MCB137L – A Feeling for the Numbers in Biology: Your Turn

Over the semester, we will do many estimates about each biological phenomenon we address. To cement these skills, MCB137L students (i.e., undegraduates) will prepare two short estimates. Your first estimate will consist of a written vignette in the style of Cell Biology by the Numbers. Some examples of interesting estimates are:

- How many proteins are in a viral capsid?
- What is the energy cost to a host cell in order to create a new virus after it has been infected?
- What is the cell-to-cell variability in the number of copies of the *lacZ* gene?
- What is the largest osmotic shock a cell can suffer without bursting?

What we're after is a vignette that feels, looks (minus having a fancy layout) and reads like a vignette from the book. Specifically, your vignette should introduce the phenomenology you're interested in making an estimate about, carefully walk the reader through the calculation, and state clear conclusions where you can compare the results of your estimates with known quantities. We encourage you to peruse through the book to get a feel for these vignettes.

Your first task is to write a short paragraph describing the estimate you're interested in writing a vignette about. Note that the objective at this point is not for you to have a finished estimate, but to have an outline of the calculation you plan to do so that we can give you feedback. Send this paragraph as an email to Hernan, Yasemin and Yovan by the homework due date. **The estimate vignette itself will be due on Thursday 3/3.**

1b. MCB237L – Physical Biology X

Graduate students (and motivated undergrads!) will develop a project involving physical modeling during the semester in what we call "Physical Biology X". The idea is to choose a biological phenomenon that you're interesting in developing a model for. This could be based on your own data, or on some paper that you find interesting (whether it already contained a model or not). For example, you could model the repressilator by Elowitz et al., develop a model of the competition between the Omicron and Delta variants of COVID-19, or learn how to use finite element simulations to calculate the diffusion and degradation of the Bicoid protein throughout a realistic fly embryo.

Your first task is to write a short paragraph describing the project you want to undertake. The idea here is for this to give you an opportunity to discuss the feasibility of the estimate with Hernan, Yasemin and Yovan. Please, send them your paragraph over email by the homework due date. **You will present your work during a poster session we will have on the last day of classes on 4/28, which will include a little celebration.**

2. DNA replication rates.

Do problem 3.3 of PBoC2 shown in Figure 1. However, as you do this problem, please come at it a few different ways. First, when estimating how much of the full fly genome is shown in the figure, account for the fact that the DNA is compacted by nucleosomes. Second, given that the entire fly genome has been claimed to have ≈ 6000 origins of replication, figure out the mean spacing between such origins and use that estimate as the basis of your own independent estimate of the replication time for the *Drosophila* genome.

3. Growth Curves and the Logistic Equation

Much of our understanding of the natural world is couched in the language of the subject now known as "dynamical systems." In a nutshell, the idea is to write down equations that tell us how some variable(s) of interest change in time. Often, this ends up being written in the form of coupled differential equations. Perhaps the most important and simplest of such dynamical systems is the law of exponential growth (or decay), relevant to thinking about the early stages of growth of a culture of cells, for example. In this problem, you are going to revisit the discussion I give there by solving for the dynamics of a population of bacterial cells both analytically and numerically.

In class, we discussed the exponential growth equation. This equation has been the basis for the study of microbiology for years (read, for example, F. Neidhardt, *Bacterial Growth: Constant Obsession with dN/dt*, J of Bacteriology 181:7405 (1999) provided on the course website). If the number of cells is given by N and the growth rate is r, then this equation takes the form

$$\frac{dN}{dt} = rN. \tag{1}$$

We solved this equation in a variety of ways, both numerically and analytically, and found a solution given by

$$N(t) = N_0 e^{rt}, \tag{2}$$

• 3.3 DNA replication rates

Assuming that Figure 3.35 is a representative sample of the replication process:

(a) Estimate the fraction of the total fly genome shown in the micrograph. Note that the fly genome is about 1.8×10^8 nucleotide pairs in size.

(b) Estimate the number of DNA polymerase molecules in a eukaryotic cell like this one from the fly *D. melanogaster*.

(c) There are eight forks in the micrograph. Estimate the lengths of the DNA strands between replication forks 4 and 5, counting up from the bottom of the figure. If a replication fork moves at a speed of roughly 40 bp/s, how long will it take for forks 4 and 5 to collide?

(d) Given the mean spacing of the bubbles, estimate how long it will take to replicate the entire fly genome.

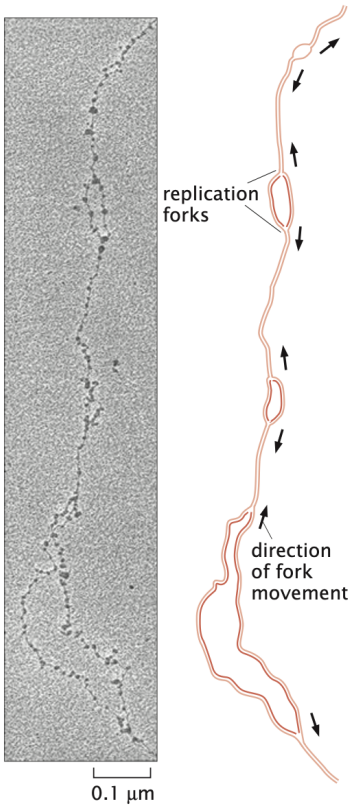

Figure 3.35: Replication forks in *D. melanogaster*. Replication forks move away in both directions from replication origins. (Electron micrograph courtesy of Victoria Foe. Adapted from B. Alberts et al., Molecular Biology of the Cell, 5th ed. Garland Science, 2008.)

Figure 1: Problem 3.3 from PBoC2.

where N_0 is the number of cells at $t = 0$.

(a) Of course, the solution shown above cannot be correct forever. For fast-growing *E. coli* estimate how long it would take for a single cell to produce enough progeny to cover the whole surface of the Earth.

A more realistic scenario is to account for the fact that, sooner or later, bacteria will run out of resources and halt their growth. For example, a liquid bacterial culture will saturate at a density of about 10^9 cells/ml. To account for these limited resources, we introduce a growth rate that depends on the number of cells, r_{new}

$$r_{new} = r\left(1 - \frac{N}{K}\right), \tag{3}$$

where K represents the maximum population size. Note that when N is very small compared to K, $r_{new} = r$ and growth is exponential. However, as N approaches K the growth rate

will decrease. Thus, we get the so-called logistic equation

$$\frac{dN}{dt} = r_{new}N = rN\left(1 - \frac{N}{K}\right). \tag{4}$$

(b) What is the number of cells at which there is no growth and the population reaches steady state? Justify how you impose steady state on the logistic equation in order to figure out this number.

(c) In class, we wrote Python code to solve Equation 1 numerically. Modify your code to now solve the logistic equation. For reasonable choices of r and K, plot number of cells as a function of time for both exponential and logistic growth.

(d) Feel free to look at section "Computational Exploration: Growth Curves and the Logistic Equation" on page 103 of PBoC2.

4. Diffusion times

Make a log-log plot of the diffusion time (in seconds) as a function of length (in μm) using Python. Plot multiple lines considering the diffusion constants for ions and for a typical protein *inside a cell*. Finally, mark a few relevant biological sizes along the x-axis such as the size of an axon, a synaptic cleft, an *E. coli* cell, and a eukaryotic nucleus.

5. Estimating the diffusion constant.

In this problem, we are going to use the observed trajectories of diffusing GFP molecules to estimate the diffusion coefficient.

(a) Conventional microscopy to observe individual fluorescent proteins moving freely in cytoplasm won't work. In this part of the problem, we are going to work out why. During a traditional experiment, the microscope shutter is open during some time interval of order 10s to 100s of milliseconds. By assuming a diffusion constant of 10 $\mu m^2/s$, work out how far the fluorescent protein will move during the time that the shutter is open and compare that distance to the size of the cell itself and comment on how this limits our ability to measure the diffusion constant. Perform the estimate a second time, this time using the 0.3 ms exposure time shown in Figure 2(A).

(b) Using the trajectories shown in Figure 3 and our simple rule of thumb that $t_{diffusion} = L^2/D$ to estimate the diffusion constant for GFP. Explain your reasoning carefully.

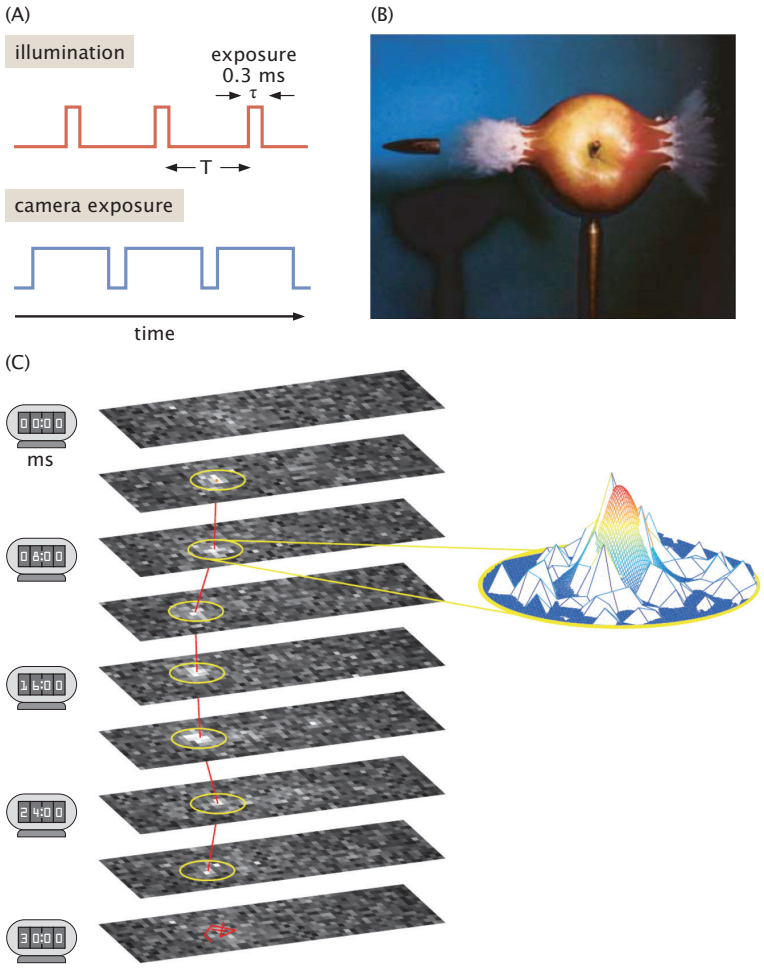

Figure 2: Stroboscopic illumination to capture fast protein dynamics. (A) By only illuminating a sample for a small fraction of the exposure time of a camera, it is possible to capture phenomena that would otherwise be blurred out. (B) A classic photo from MIT legend Harold Edgerton who pioneered stroboscopic photography for science and fun.Capturing the piercing of a bullet through an apple using stroboscopic illumination. (C) Measuring the position of an individual GFP molecule inside *E. coli*. (A, adapted from Harold and Esther Edgerton Foundation, 2006, courtesy of Palm Press, Inc.)

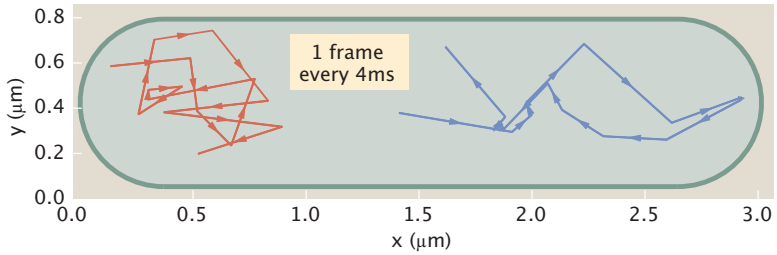

Figure 3: Time series showing positions of diffusing GFP molecules at different times. The red and blue traces correspond to different molecules. The lines are a guide to the eye. Adapted from BP English *et al.* Proc. Nat. Acad. Sci., 108:E365-E373, 2011.