MCB137L/237L: Physical Biology of the Cell
Spring 2020
Homework 1: Biological Numeracy
(Due 1/30/20 at 3:30pm)

Hernan G. Garcia

"The greats weren't great because at birth they could paint. The greats were great cause they paint a lot." - Macklemore

Homeworks in MCB137/237

Whether it is in the context of professional sports, art, or science, it pays to practice. The idea of the homeworks during our course is to give you a venue to get your 10,000 hours of calculations and estimates in as a means to become proficient in the mathematical and physical modeling of living systems. Sometimes, the problems will ask you to redo a derivation we did in class in a new way, and sometimes they will propose a whole new biological phenomenon to attack. Regardless, if you spend more than five hours on a homework set, it means that you should come to office hours. Make sure to start working on your problems early on!

The objective of this homework is to get a feeling for the numbers in whatever problem you're considering in biology. Just like you always need to check the units in your calculations, a more subtle sanity check of your theoretical results stems from having some expectation about the order of magnitude you will obtain.

When doing street fighting estimates, the goal is to do simple arithmetic of the kind that all numbers are 1, few or 10. few $\times$ few $= 10$, etc. Please do not provide estimates with multiple "significant" digits that are meaningless. Be thoughtful about what you know and what you don't know. You may use the Bionumbers website http://bionumbers.hms.harvard.edu/ to find key numbers (examples are masses of amino acids (BNID 104877) and nucleotides (BNID 103828), the speed of the ribosome (BNID 100059), etc.), but please provide a citation to the Bionumber of interest as shown above. However, for many of these problems the essence of things is to do simple estimates, not to look quantities up.

Sometimes, the problems will be drawn directly from the 2nd edition of Physical Biology of the Cell (PBoC or PBoC2). In that case, I'll make the effort to scan the problems and include them as a figure. However, some of those problems might refer to information inside the book, which I will not scan. As a result, I highly recommend that you just get the book.

Homework submission: Gradescope will be used to submit and grade your homework. We will create two submissions for weekly homework (one for written pdf, the other one for submitting zip file for your code). When you have to write Python code in order to make plots, you don't need to include your actual code in the pdf document you submit. However, you need to submit all the codes you used to generate the plots in a separate zip file. All plots you generate need to have axes and lines that are clearly labeled. Please submit both pdf and original code before the deadline.

Finally, remember to write each problem on a different piece of paper so that you can upload them independently to Gradescope. This will make it easier for us to grade them.

How to join Gradescope:

1. Go to website: https://www.gradescope.com
2. Create an account
3. Add class with entry code: **M8YRBJ**
4. Please update your "Student ID" in the account settings

1 Sizing up the Central Valley

California's Central Valley is one of the most potent agricultural regions in the world. In this problem, you are going to evaluate many of the key factors associated with its enormous productivity without any data aside from a single satellite image of the region as shown in Figure 1. Note that the key point here (and what you will be graded for if you care about such things) is the logical flow of your estimates, not the particular numerical values you found.

(a) Water usage. Using what you know about watering and the growth of plants, make an estimate of the amount of water used to irrigate the agriculture of the Central Valley.

(b) Nitrogen usage. Since the beginning of the twentieth century, we have doubled the number of occupants that can be fed on Earth as a result of the Haber-Bosch process and the synthetic fixation of nitrogen. In this part of the problem, begin by estimating the number of kilograms of biomass per square meter that is produced per year. From that number, figure out how many kilograms of nitrogen are contained per square meter of biomass. Then, make an estimate of how much fertilizer is used for each square meter and hence for the entirety of the Central Valley.

(c) Pesticide usage. Undertake an estimate similar to that in the first two parts of the problem to figure out how much pesticide is used on the Central Valley every year.

(d) Do NOT do this part until you have done parts (A) - (C). Look up some source of data on each of these three questions and compare your results to the data. Please do not redo

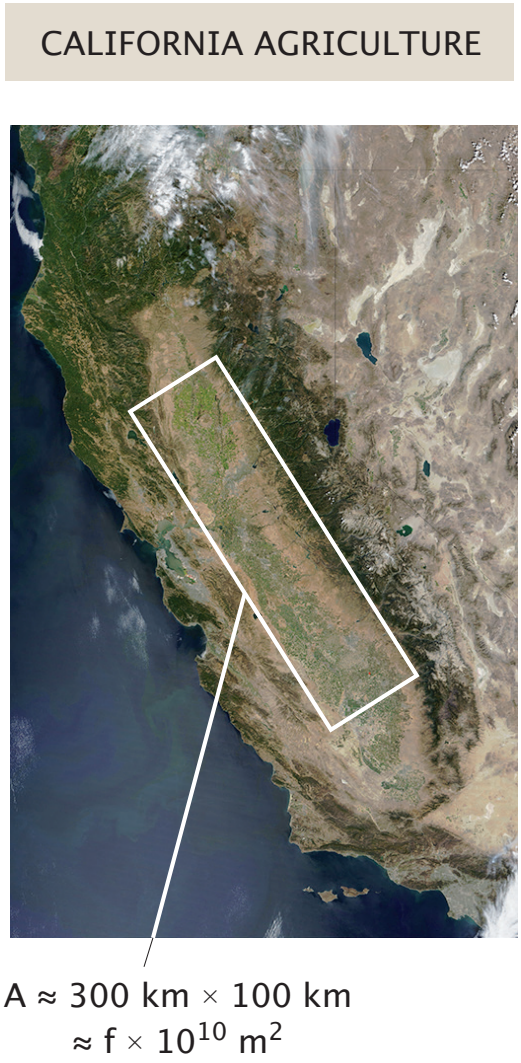

Figure 1: Satellite image of California's Central Valley.

your estimate.

2 DNA Synthesis Over Your Lifetime

Estimate the total length of DNA your body will produce over your lifetime. To make this estimate, you can first figure out which cells are the most numerous in your body by reading Sender *et al.*, Cell 164:337 (2016), which is provided on the course website. Then, find out how often these cells get renewed.

3 Building a bacterial cell

Do problems 2.5 (ingredients in minimal media) and 3.7 (sugar budget of a cell) from PBOC2. Both problems are shown in Figure 4. Together, these two problems are intended to get you thinking about the wondrous process whereby cells convert a clear liquid with simple chemical ingredients into biomass as shown in Figure 2. Amazing! You can use Figure 3 or refer to BioNumbers to look up the composition of the cell.

After working out the two problems given above, work out an estimate related to the volume of the headspace you see in Figure 2 which has oxygen available for cell growth. Specifically, if 6 O_2 molecules are consumed for every sugar, make a simple estimate of the required volume of headspace needed to sustain cell growth. Note that our estimate about O_2 usage is crude and sloppy. To really do this carefully, we need to acknowledge the use of glucose both in providing building materials (i.e. carbon skeletons) as well as the energy needed to synthesize a cell. The estimate we do here is intended to give an impression of the magnitudes, and specifically to get a sense of the aeration requirements when we do a liquid culture growth procedure.

Do problems 2.5 and 3.7 from PBOC2 (shown in Figure 4). Together, these two problems are intended to get you thinking about the wondrous process whereby a clear liquid with simple chemical ingredients is converted into biomass.

4 Street fighting your way to the ribosome density

One of the most important molecular assemblies in the cell is the ribosome. The number of ribosomes per cell dictates how fast cells can grow. *E. coli* growing with a division time of 24 minutes have 72,000 ribosomes per cell, and slow growing *E. coli* with a division time of 100 minutes have a factor of ten fewer ribosomes with a count of $\approx$ 6800 ribosomes. In this problem, we will use our street fighting skills to explore the ribosomal density in another organism as shown in Figure 5, and then see how well our results from the electron microscopy study square with the numbers quoted above. By examining the figure, make an estimate of the number of ribosomes per μm^3 and compare that result to the numbers quoted for *E. coli* above.

Figure 2: Growth of *E. coli* in rich media. The tube on the left shows roughly 5 mL of growth media just after inoculation. The tube on the right shows such media after saturation due to exponential cell growth and division.

5 RNA vs. Protein

Using the kind of estimates we have talked about in class, give a simple characterization of the relative sizes of mRNAs and the proteins they code for. Specifically, first comment on the mean mass of amino acids and nucleotides as well as their typical physical sizes. Use both of these metrics as a way to provide a rough sense of how both mass and physical dimensions for proteins vs. the mRNAs that code for them.

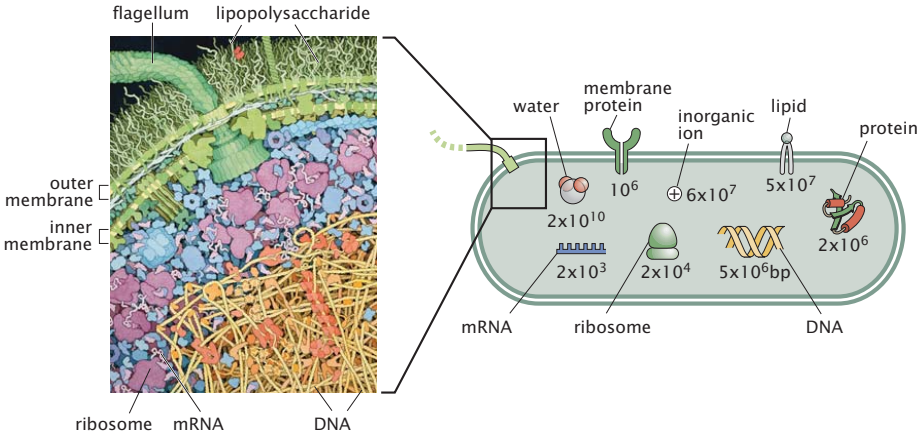

Figure 3: Molecular contents of the bacterium *E. coli*. The illustration on the left shows the crowded cytoplasm of the bacterial cell. The cartoon on the right shows an order-of-magnitude molecular census of the *E. coli* bacterium with the approximate number of different molecules in *E. coli*. (Illustration of the cellular interior courtesy of D. Goodsell.)

• **2.5 Minimal media and *E. coli***

Minimal growth medium for bacteria such as *E. coli* includes various salts with characteristic concentrations in the mM range and a carbon source. The carbon source is typically glucose and it is used at 0.5% (a concentration of 0.5 g/100 mL). For nitrogen, minimal medium contains ammonium chloride (NH_4Cl) with a concentration of 0.1 g/100 mL.

(a) Make an estimate of the number of carbon atoms it takes to make up the macromolecular contents of a bacterium such as *E. coli*. Similarly, make an estimate of the number of nitrogens it takes to make up the macromolecular contents of a bacterium? What about phosphate?

(b) How many cells can be grown in a 5 mL culture using minimal medium before the medium exhausts the carbon? How many cells can be grown in a 5 mL culture using minimal medium before the medium exhausts the nitrogen? Note that this estimate will be flawed because it neglects the *energy* cost of synthesizing the macromolecules of the cell. These shortcomings will be addressed in Chapter 5.

• **3.7 The sugar budget in minimal medium**

In rapidly dividing bacteria, the cell can divide in times as short as 1200 s. Make a careful estimate of the number of sugars (glucose) needed to provide the carbon for constructing the macromolecules of the cell during one cell cycle of a bacterium. Use this result to work out the number of carbon atoms that need to be taken into the cell each second to sustain this growth rate.

Figure 4: Problems on building a bacterial cell from PBoC2.

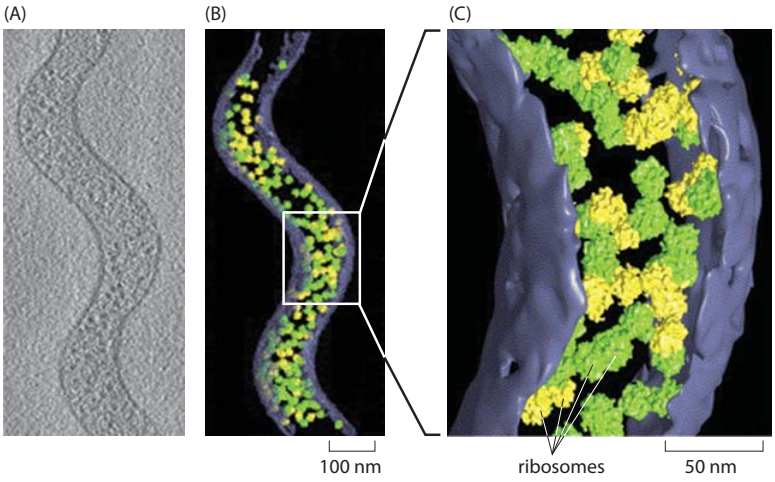

Figure 5: Cryo EM study of a bacterial cell. These images are of the tiny bacterium, *Spiroplasma melliferum*. Using algorithms for pattern recognition and classification, components of the cell such as ribosomes were localized and counted. (A) Single cryo-electron microscopy image. (B) 3D reconstruction showing the ribosomes that were identified. Ribosomes labeled in green were identified with high fidelity while those labeled in yellow were identified with intermediate fidelity. (C) Close up view that you should use to make your count. Adapted from JO Ortiz *et al.*, J. Struct. Biol. 156, 334-341 (2006).