Biol 213 Genetics: Wednesday, November 15, 2000
Mutagenesis (Part I)

I think it is fair to say that this past exam was a harrowing experience for most if not all of us (and I choose my pronouns with care). Some despaired that they had worked hard but had nothing to show for it. While respecting the sentiment, I must emphatically disagree. Hard work always pays off, though not necessarily at the moment you anticipate. Isometric exercises may not budge the wall, but the muscles get stronger. To my mind, last night was barely a pop quiz. The true test will not occur until, say, ten years from now when you are called upon to find your way through a confusing problem, perhaps this time with a life in your care. The more you fight through the thickets, the better equipped you are in the future to do so with success

Nonetheless it is discouraging to work hard and reach far and still encounter problems that lay beyond your grasp. No doubt the exam was bad psychology -- after three weeks of confusing material you needed a boost -- but try for a moment to see it not as a judgment but as a journey. Look objectively at the problems you attacked and for a moment let go of concerns over what you did not understand. Compare these questions with those you may have answered a year ago. Put them side by side -- aren't you struck by how far you've traveled? I don't know that any of us ever finishes the journey. What matters is how you travel.

Outline (topics in italics to be covered on Friday)

I. Introduction

II. The clonal nature of mutation (pp.463-466)

III. Spectrum of spontaneous mutations

III. Types of base substitutions
- A. Spontaneous base substitutions (pp.455-456)
- *B. Environmentally-induced base substitutions (pp. 469-471)*

IV. Insertion/deletion
- *A. Insertion/deletion at the lacI hotspot*
- *B. Insertion as a cause of human disease (pp.459-460)*

V. Repair of mutation (pp.472-478)
- *A. Bypass replication*
- *B. Excision repair*
- *C. Mismatch repair*

I. Introduction

Take a look at the picture to the right (Fig. 1). ***That's you, if mutation had halted a couple of billion years ago.*** You may well conclude that mutation is not necessarily such a bad thing, despite the bad press it generally gets. While change has paid off for us over evolutionary time, for viruses it is a way of life. HIV has proved a formidable foe because it mutates so rapidly, thereby sidestepping obstacles that we put before it.

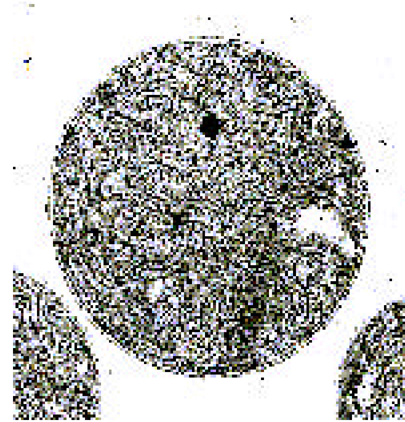

Fig. 1. Cell of the cyanobacterium *Anabaena.* Cyanobacterial remains have been found in fossils over 3.5 billion years old.

The key to understanding mutations is the three R's: Rare, Random, and Recursive. I've been harping on rare all semester, so that's not news, and we've used the random nature of mutations as justification for multiplying their probabilities. But if mutations are so rare and random, then how come a person can end up with a large tumor, every cell of which is multiply mutated? To understand this, we need to understand recursion, the idea that one mutation can build on another over many cellular generations. Another way of describing this idea is by the word **clonal**: mutations are propagated by the growth of cells that possess them.

II. The clonal nature of mutation (pp.463-466)

It wasn't too long ago that mutation was considered in very different terms than it is today and for intuitively plausible reasons. Consider the common laboratory situation that a culture of bacteria is exposed to some harsh condition, say an antibiotic. For a long while, growth ceases, but quite often growth resumes, with the culture now insensitive to the antibiotic's effects. Isn't it reasonable to presume that the stress of the antibiotic induced the bacteria to retool and find a way to overcome the threat?

The modern way of looking at the situation is that the antibiotic merely killed off the vast majority of bacteria that were sensitive to it, leaving the insensitive minority to take over the flask. The critical difference between these two views is that mutation in one case is caused by the challenge while in the other case mutants exist prior to the challenge, which merely selects for them. Which is it?

In 1942, Salvadore Luria and Max Delbrück found a way to put the question to the test. Bacteriophage are a potent threat to bacteria, but from your own experience you know that a small minority of *E. coli* are able to survive the challenge and make colonies, even when exposed to excess phage T4. How can you tell if these mutant *E. coli* are selected by exposure to T4 or are caused by the phage?

SQ1: Suppose that you add 10^9 phage to 10^8 *E. coli* cells, and you find that about 100 cells survive to make viable colonies. Is this scenario consistent with the idea that the phage induces mutation to phage resistance? Is it consistent with the

idea that a small number of phage-resistant bacteria exist in the culture before exposure to phage?

Luria and Delbrück realized that though either model would lead to the same result on average, but it was still possible to distinguish them. They divided a culture of *E. coli* amongst twenty tubes and after allowing a period of growth, challenged each with bacteriophage. If the bacteriophage causes the mutation, then it made no difference if the culture were grown in twenty tubes or pooled together in a single flask. Either way, you would expect the same number of phage-resistant colonies per plate. The left hand side of Fig. 17-6 illustrates this scenario. Each sample from the flask onto plates seeded with phage put about the same number of phage-resistant cells. The variation in the number resulted just from the variation in sampling small numbers.

If variation is due solely to sampling error (i.e., there is no significant variation from the pipetting), then it is possible to calculate the variation from the average number of colonies according to the equation:

standard deviation = Square root(mean value)

This is an extremely useful relationship to take with you wherever you go (see Problem Set 10).

The situation is quite different if mutations occur at random before exposure of the culture to bacteriophage. Suppose that a mutation arises by chance in a tube just before the phage is added. In that case, a single phage-resistant colony will appear on the plate. Alternatively, suppose that the mutation arose four generations prior to the addition of phage. Then that single mutant cell will have the chance to double four times, leading to sixteen phage-resistant colonies. A single mutation event can have very different consequences for the number of resistant colonies, depending on when in growth the mutation occurs. Thus twenty different replicates of the experiment can result in wildly different numbers of resistant colonies, as illustrated in the right hand side of Fig. 17-6. Note that the average number of colonies in both cases is about the same.

SQ2: What would have been the result in the experiment shown in Fig. 17-6 if the twenty tubes had been pooled, mixed, and then redistributed to twenty new tubes immediately prior to plating on phage-seeded medium?

The Luria-Delbrück experiment was very influential in tearing down the last bastion of Lamarckianism. Just as giraffes don't pass on longer necks to their progeny by stretching them while they eat, neither do bacteria pass on phage resistance caused by exposure to phage.

The power of the experiment was in the mathematical analysis of variability, an arena in which few biologists were comfortable. It was therefore a welcome event when ten years later Lederberg and Lederberg pitted the same opposing hypotheses of random vs induced mutation by means of a qualitative experiment (Fig. 17-7). As with the Luria-Delbrück experiment, *E. coli* were allowed to grow without selection under

conditions of nonmixing. This time, the growth occurred on a plate rather than separate culture tubes.

SQ3: Suppose that phage resistance is induced by the presence of phage. How many mutants would you expect to exist on the plate prior to exposure to phage?

SQ4: Suppose that phage resistance arises at random in cells that continue to grow on the plate and thus propagate the mutation. If you find a mutant at a particular spot on the plate, is it likely or unlikely that you could find more at precisely the same spot?

The Lederbergs could have continued in a way that paralleled the approach of Luria and Delbrück, by sampling different spots on the plate and measuring the variability in the number of phage-resistant mutants obtained. They instead took a different approach, sampling the entire plate twice, by replica plating (Fig. 17-7). If mutants arose on the replica plates at the same positions, then there must have been multiple mutant cells a that position on the parent plate.

SQ5: What would be the expected results from the Lederbergs' experiment if phage induced phage resistance in *E. coli.*

III. The spectrum of spontaneous mutations

What are mutations? At root, they're changes in the DNA sequence in a gene or in the sequence determining the regulation of a gene. Suppose I tried to consider all the possible ways that a sequence could be changed. Surely I'd be biased towards mutations that make sense to me. Twenty-five years ago, for example, no one would have thought to look for mutations in introns -- no one knew introns existed! Better, let me simply ask an organism: Tell me about your mutations. Let's look at a couple of hundred random mutations and look to see what they are.

Easier said than done, for mutations are exceedingly rare. How can we collect hundreds of mutations and then analyze them? We need a way of selecting mutations. It won't do to look at *E. coli* one by one and wait for a mutation to pop up, especially when a gene is mutated at a frequency of 1 in a million cells. By "selection" I mean a procedure that kills off every cell that doesn't have a desired mutation. In this way, even if only one in a million cells survive the selection, they're all mutants.

Can you think of a well-studied gene of *E. coli* where you can select for loss of function? Yes, the *lac* operon -- again. Recall that Jacob and Monod obtained *lacI* mutants by growing *E. coli* on PGal, because PGal can be metabolized by ß-galactosidase but it cannot induce the *lac* operon. It can therefore support growth only when the repressor is dead. Why does this damn operon keep popping up all over genetics? It really isn't that they're remarkable genes, it's just that people have been studying *lac* for more than 50 years, and -- here's the take-home lesson -- the more you study something, the more it can illuminate everything else you do.

So now we have an experiment (Fig. 2).

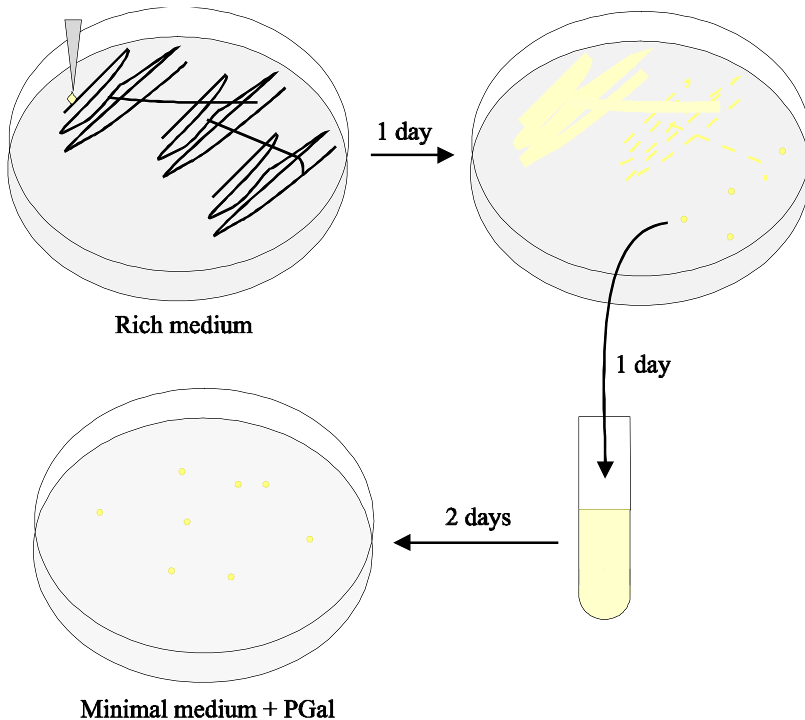

Fig. 2. Procedure to select for *lacI* mutants. A drop of *E. coli* culture was placed on a petri dish with rich medium and streaked out, so that well-isolated colonies would appear the next day. One of the colonies was grown up in liquid medium. The culture was spread on a minimal medium plate in which Pgal was the sole source of carbon. Cells with mutant *lacI* genes grew at the expense of Pgal and produced viable colonies. The colonies were grown up, and the *lacI* gene was sequenced from isolated DNA.

SQ6. Why begin the experiment with a single colony, arising from a single cell? Why not start by regrowing *E. coli* from an old culture? Consider that even a small volume of old culture (say, 0.01 ml) necessary to start a new culture would contain (0.01 ml) x (2 x 10^9) bacteria. If this hint isn't enough, revisit this question after viewing the results of the experiment.

SQ7. Why plate the culture on PGal?

SQ8. What kind of mutations might you expect to see in a gene? Consider what kinds of changes in the DNA would lead to the phenotype you're looking for.

Table 1 shows an analysis of 176 *E. coli* mutants defective in *lacI*. Only 11% of the observed mutations were due to base substitutions, and single base frameshifts (like those studied by Crick et al) were even more rare, only 3% of the total. Fully 86% of the mutations were by mechanisms we haven't dealt with thus far.

TABLE 1: Spectrum of spontaneous mutations in *lacI*[1]

Class of mutation	Occurrences	Frequency[2] x 10^{-7}
Base substitutions	20 (11%)	2.2
Single base frameshifts	5 (3%)	0.7
Insertion elements	7 (4%)	0.8
Duplications (total)	98 (58%)	11.3
hotspot[3]	*95 (54%)*	*11.0*
Deletions (total)	44 (26%)	5.0
hotspot[3]	*22 (13%)*	*2.5*
Unknown	2 (1.1%)	0.2
TOTAL	176 (100%)	20.2

[1]Mutations were selected by growth on Pgal, a galactoside that is broken down by ß-galactosidase but does not efficiently bind to the Lac repressor. Data from Schaaper et al (1986) J Mol Biol 189:273-284.

[2]Frequency represents the number of mutations recovered per cell.

[3]Mutations at "hotspots" occurred specifically at pUR3 coordinate 2778.

Ask an unbiased question, get an unexpected response.

SQ9. What is this hotspot that got all the action? Look up in your lab manual the sequence of the region in *lacI* (pUR3 coordinate 2778). See anything unusual?

Clearly, the biggest surprise from this and similar experiments is that mutations are far from random. 67% of the mutations recovered in *lacI* occurred at a single site. While mutations may arise at random in time, they certainly don't arise at random in space. We'll take a look at that hotspot and see what it can tell us about common mutations in humans as well, but first we'll consider base substitutions, a class of mutation that will be particularly important when we look at environmentally induced mutagenesis.

IV. Mechanisms of Base Substitutions (pp.584-586)

III.A. Spontaneous base substitution

11% of all mutations in *lacI* were base substitutions (Table 1). One of the most common ways for base substitutions to arise is by spontaneous tautomerization of one of the four bases that alters its base pairing properties (Fig 3[1]). It is an unfortunate fact of nature that rare but naturally occurring forms of guanine pair with thymine, adenine with cytosine, cytosine with adenine, and thymine with guanine. If any of these forms arise, there is the possibility of a base pair changing as illustrated below (Fig. 4):

The properties of tautomerization allow for only certain kinds of mutations: A→G, G→A, T→C, and C→T. These kind of mutations, where a purine mutates to the other purine or a pyrimidine mutates to another pyrimidine, have been given the name "transitions", while those not resulting from tautomerization, A→C or T, G→C or T, C→A or G, and T→A or G, are called "transversions".

You might expect, then that an examination of spontaneous mutations would reveal a preponderance of transitions over transversions. Let's take a look (Table 2).

SQ10: What fraction of the base substitutions found in *lacI* are transitions? What fraction are transversions?

Why were our expectations ignored? First, the experiment was biased in the mutations that it detected in *lacI*. Examine the genetic code shown on p.327 and in particular the

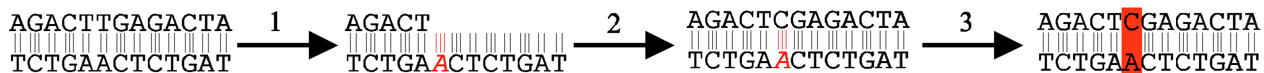

Fig. 4: Base substutation by tautomerization. AT and GC base pairings are shown by two or three lines, respectively, to represent appropriate number of hydrogen bonds. (1) During replication, an adenine (shown in red) spontaneously tautomerizes to a form that can base pair with cytosine. (2) When the other strand is replicated, a cytosine is placed opposite the altered adenine. (3) When the altered adenine reverts to its normal form,an AC mismatch results. When the two strands are replicated again, the upper strand will lead to a GC pair at the affected position, while the bottom strand will lead to the wild-type AT pair.

[1] This figure is supplied separately, owing to copyright restrictions. It is available only to those in the UR community.

codon UUG (or TTG in DNA). A transition at the third position (TTG → TTA) or a transition in the first position (TTG → CTG) does not change the amino acid! In contrast, all transversions affecting the codon cause an amino acid change.

SQ11. Did I choose a special case? In general, do transitions more often than transversions lead to amino acid changes?

The foregoing discussion reminds us that base subsitutions can have different levels of effect (see also Table 17-1 in the text):

Silent mutation	No amino acid change
Conservative mutation	Change from one amino acid to a similar amino acid
Missense mutation	Change from one amino acid to another
Nonsense mutation	Change from one amino acid codon to a stop codon

There is a second reason why transitions may be underrepresented in the collection of *lacI* mutants. Both prokaryotes and eukaryotes stop DNA replication the moment they discover that there's something wrong with their DNA. This gives them time to repair the damage. In humans, if this defense mechanism is lost for some reason, cancer is often the result. On Friday we will examine the various ways cells try to repair damaged DNA, and on Monday we will examine what happens when these attempts fail.

SQ12. The CAG codon (encoding glutamine) more commonly mutates to TAG (stop) than to AAG (lysine). Why?

Table 2: Spectrum of spontaneous *lacI* mutations[1]

Position[2]	Change	Occurrences	Amino acid change
54	AT to GC	1	Thr to Ala
112	TA to AT	1	Val to Glu
121	CG to AT	1	Ala to Glu
129	GC to TA	1	Glu to *TAG*
154	TA to AT	1	Val to Glu
184	TA to AT	1	Leu to *TAG*
193	GC to AT	1	Gly to Asp
264	CG to TA	1	Gln to *TAA*
390	CG to TA	1	Gln to *TAG*
553	TA to GC	1	Leu to *TGA*
748	CG to AT	1	Ala to Glu
816	TA to GC	1	Tyr to Asp
856	TA to GC	1	Ser to *TGA*
897	AT to CG	2	Ser to Arg
lacO[3]	TA to CG	5	O^C

[1]Data taken from Schaaper et al (1986) J Mol Biol 189:273-284
[2]Beginning of *lacI* is coordinate 1. Add 2186 to get pUR3 coordinates.
[3]pUR3 coordinate 3358, about in the middle of the region where the repressor binds.