

Recombinant DNA—Libraries and Forensics

A. Library construction

Recall that earlier in the semester we considered the *Cys4* yeast gene that encodes a CBS protein. Recall also that yeast mutant in *Cys4* are cysteine auxotrophs, while humans mutant in a homologous gene have a host of medical problems from mental retardation to heart disease and ocular degeneration.

Today we will talk about finding the yeast gene responsible. In Section 22 we will turn to the question of using a yeast system to diagnose this particular human disease.

Suppose we wanted to find the particular gene responsible for the cysteine auxotrophy phenotype. We decide to use cloning by complementation, and we use one of our cysteine auxotroph *cys4⁻* mutant strains as a host strain.

1. What is a library?

A collection of DNA molecules, derived from restriction fragments that have been cloned in vectors, that includes all or part of the genetic material of an organism.

2. What is a library used for?

We are talking about using a library to clone by complementation. There are other possible uses, such as cloning by hybridization – finding a vector that contains a fragment of the genome that is complementary to the nucleic acid probe.

3. What is a common principle between complementation by mating that we did in Genetics and identification of a gene using genomic library?

In both cases we seek to create an organism that is diploid in the gene/genes of interest to determine whether the presence of the second copy of the gene would complement the deficiency. In complementation by mating, we create a diploid organism. In the course of cloning by complementation using a library we create a partial diploid – an organism diploid for the gene or genes carried by the particular vector that was taken up by the cell.

We need a vector and a source of genomic DNA to construct a library.

4. What must the genotype at the *Cys4* locus be of the strain of yeast that we use as the source of genomic DNA? What phenotype must it have? Why?

*The strain we use must be *Cys4⁺*. Because we are cloning by complementation, and our host strain is *cys4⁻*, the library must contain a *Cys4⁺* allele. Otherwise, no complementation is possible. The phenotype of the strain used is not restrictive. We can use cysteine auxotroph strains, as long as they are mutant in genes other than *Cys4*.*

5. What features should the vector we are using have?

Origin of replication, restriction site, and a selectable marker (for example, Kan resistance)

6. If we ran a gel of the vector and genomic DNA after the restriction digestion, how many bands would you expect to see in the vector digest lanes? What does your answer depend on?

We would expect as many bands in the vector lane as there are restriction sites in our original vector for our restriction enzyme of choice. However, depending on the amount of time a digestion is allowed to proceed, we may see the number of bands corresponding to the number of restriction sites for our enzyme, or more, corresponding to various partial digests. In general, in this course you will only see the gels where DNA has had a chance to be fully digested.

7. How many cut sites for our restriction enzyme do you expect in the yeast genome? How many bands would that produce?

Many sites distributed randomly throughout the genome. In fact, so many sites that we should see a smear on the gel because we would not be able to distinguish individual bands. Since a cut site is a particular sequence of six base pairs, we would expect to see on average one such site per 4^6 bases. Of course, the sequence will not show up at regular intervals, so we can expect to see a smear on the gel.

8. What are the next three steps in library creation? What is the goal of each step?

Step	Goal
Ligation	To create closed vectors carrying a fragment of yeast genomic DNA each
Transformation	To introduce the vectors created in the previous step into a culture of yeast. The strain we choose should allow the use of selectable marker present on the vector.
Selection	To find the yeast colonies resulting from taking up the vectors containing our gene of interest.

9. If we plated our transformed culture on Complete+Kan media, what would each colony on the plates represent?

Each colony would represent an original yeast cell that took up a vector from our library.

10. What media should we plate on to find the cells that took up a vector from the library that contains a wild-type Cys4 gene?

We should plate on Cysteine dropout +Kan media. Kan will select for taking up the vector, while the absence of cysteine in the media will select for those cells that can now make their own cysteine, i.e. for those cells where complementation occurred.

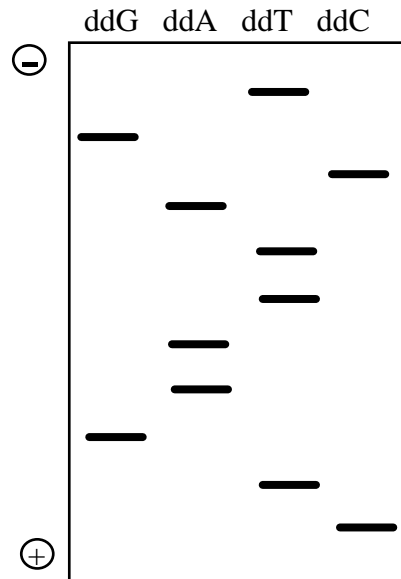
B. Application—Solve a mystery

BARNEY AND THE BEAVER ARE GONE!! Help us solve the mystery so Professor Walker can continue to bring his trusted visual aids to 7.014.

FACTS:

- Barney and the beaver were seen by all in lecture on Wednesday, February 2.
- The last time Barney and the beaver were seen was at a staff meeting in Professor Walker's office immediately after that lecture.
- Melissa, Luke, Cindy, Eric, Julia, and Professor Walker were all present at the meeting.
- One of these staff members (not Professor Walker) took the props.

The techniques of dideoxy DNA sequencing and the polymerase chain reaction (PCR) were used to solve the crime. Cells left on the ransom note were isolated and the region containing gene K amplified by PCR. The DNA sequence of the PCR products was then determined. Shown in the box below are the results from a sequencing gel of a region of gene K.



1. Write the DNA sequence for this region of gene K as shown in the sequencing gel above. Indicate the 5' and 3' ends of the DNA sequence.

5' CTGAATTACGT 3'

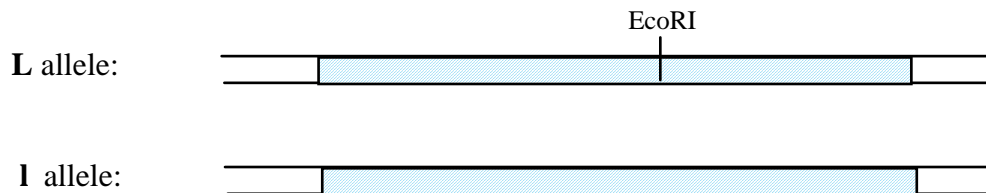
2. One of the suspects, Cindy, has the following ds DNA sequence for this same region in both copies of gene K:

5' CTGAAGTACGT 3'
3' GACTTCATGCA 5'

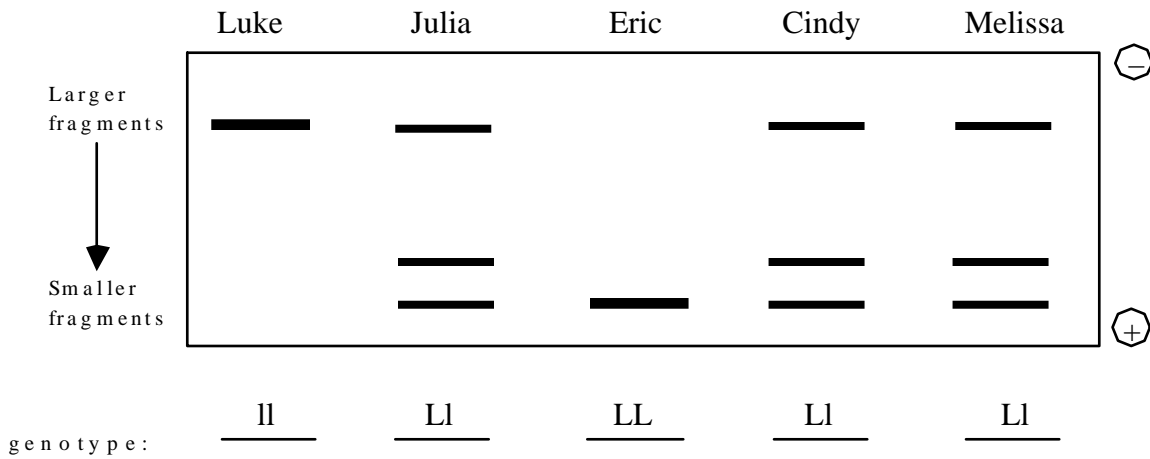
Is Cindy still considered a suspect in the case? Justify your answer.

No, Cindy's DNA sequence for gene K is different than that found on the note.

Analysis of a second site, the gene L locus (represented by the shaded box below), is useful for forensics. The only difference between the two alleles at this locus is that the **I** allele lacks an internal EcoRI restriction site present in the **L** allele.



PCR analysis of cells taken from the note indicated that it was handled only by a **LI** individual. Based on this, the gene L locus of the suspects was amplified by PCR, the PCR products were digested with EcoRI, and the resulting fragments were separated by agarose gel electrophoresis as shown below:

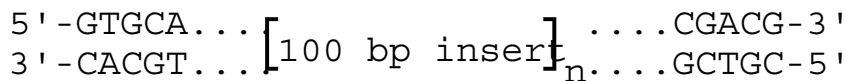


- In the spaces provided above, indicate the genotypes at the gene L locus for these suspects.
- Based on the PCR analysis of the gene L locus, which individual(s) can be excluded as suspects in the crime? Justify your answer.
Luke and Eric. Their DNA samples at the gene L locus do not match the sample found on the note.

At chromosomal position Z, it is common to find one or more copies of a 100 bp insert. Fortunately, most of the suspects have a different number of inserts at this position. We even tested Professor Walker to be sure:

- Eric has no inserts.
- Luke has one insert.
- Melissa has two inserts.
- Cindy has two inserts.
- Julia has three inserts.
- Professor Walker has four inserts.

A partial DNA sequence of position Z is given below. The length of position Z without an insert is 2000 bp.

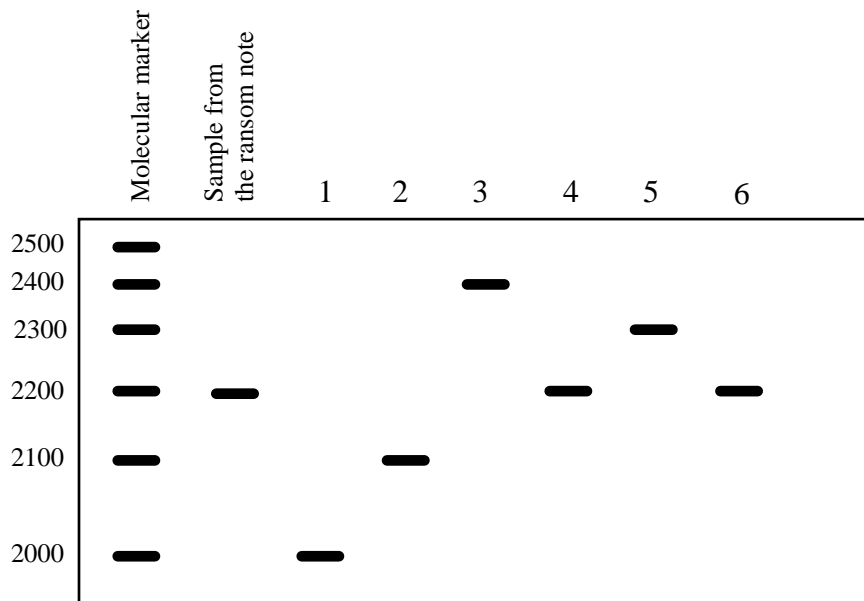


- To determine who took Barney and the beaver, design a PCR-based strategy to analyze position Z, including the sequences and orientation (5' and 3' ends) of two 5-base primers necessary to PCR amplify this region.
You design the following two primers to PCR amplify position Z.



You obtain DNA from cell samples taken from all suspects and from the scene of the crime. You PCR amplify position Z from all samples and obtain PCR products. You size separate the PCR products by gel electrophoresis. The size of the PCR product amplified from position Z from the crime scene will match the pattern for one of the suspects.

6. The following pattern was seen on the gel used to separate the PCR products from the above experiment.



i) Which lane corresponds to which individual?

1: *Eric*

2: *Luke*

3: *Professor Walker*

4: *Melissa or Cindy*

5: *Julia*

6: *Melissa or Cindy*

ii) Who committed this heinous crime?

Melissa must have committed this crime. Only Cindy and Melissa match the sample taken from the note. But Cindy is no longer a suspect due to part 1.