C2005/F2401 '08 -- Lecture 17 -- Last Edited: 11/09/08 11:45 AM (Problems to do are indicated in **red bold.**)

Handouts 17A &B -- not available on the web; extra copies will be in boxes outside 744 Mudd after the pm lecture.
17A = Restriction enzymes, RFLP's, & Overview of Making a Chimeric Plasmid; 17B = Blots & A Cloning Vehicle.

Note: References to texts are to Sadava (Purves) 8th ed. & Becker 7h ed. (References to 7th ed. of Purves and 6th ed. of Becker are in parentheses if different.) Recombinant DNA is covered in different chapters in 6th ed. of Purves (as vs. 7th & 8th) & in 5th ed. of Becker (as vs. 6th & 7th).

I. Viral Life Cycle & Viral Genetics

A. Lysogenic Cycle See Sadava 13.3 (13.2) or Becker Box 18A, esp. fig. 18A-4.

1. Integration. Some viruses can become part of the host chromosome by crossing over between the viral DNA and the bacterial DNA -- the process is parallel to the way a plasmid such as the F factor joins the bacterial chromosome. (Viruses, like plasmids, can pick up bacterial genes by the reverse of this process.)

2. Lysogeny. The integrated virus can remain dormant for long periods of time. This dormant state is known as lysogeny, and a bacterium with an integrated, dormant, virus is said to be lysogenic (capable of entering the lytic cycle). What keeps the virus from making viral proteins and entering the lytic cycle? A repressor protein made by the virus itself. (This repressor protein is not allosteric; it must be destroyed to be inactivated. The degradation of repressor protein allows the virus to leave the dormant state and enter the lytic cycle.) Jacob, Lwoff, & Monod received the Nobel Prize in Physiology in 1965 for figuring out how repressors control both operons and lysogeny.

B. Retroviruses.

1. Need for Reverse Transcriptase: These are viruses (not necessarily of prokaryotes) that contain RNA in the viral particle. When the RNA enters the cell, it uses a special enzyme made by the virus, reverse transcriptase, to make a DNA copy of the RNA. (The reverse transcriptase is carried into the host inside the viral particle.) The DNA then inserts into the host chromosome and remains dormant, in much the same way as a lysogenic virus. HIV, the virus that causes AIDS, is a human retrovirus.

2. Significance of Reverse transcriptase: Reverse transcriptase obtained from retroviruses is used in the lab as an important tool to make DNA copies of RNA. (Examples will be discussed next time.) For HIV life cycle see Sadava 13.6 (13.5). The 1975 Nobel Prize in Physiology was awarded to Dulbecco, Temin & Baltimore in 1975 for the discovery of reverse transcriptase in tumor causing viruses.

C. Viral Crosses -- Complementation & Recombination can occur with viruses as well as with bacteria. If a cell is simultaneously infected with two variants (mutants) of the same virus, then crossing over and/or complementation can occur between the two viruses during the course of infection. See last lecture, V-B-3, and handout 16B for more details on complementation (& how to distinguish complementation and recombination).

For an example, see problem 11-8. (For more problems involving compl. & recomb. in viruses, see 11-10 to 11-13.)

See Becker fig. 20-17 (20-18) for recombination in viruses. (Re-assortment may also occur in the case of flu virus, which has an RNA genome segmented into 8 pieces. See CDC page for more details. For life cycle of an RNA virus, see Sadava 13.5 (13.4)

II. Restriction Enzymes

A. Introduction: The idea/problem: The existence of plasmids and bacteriophages carrying bacterial genes inspired dreams of genetic engineering. Why not make new combinations to order? That way we could make plasmids with useful genes and add them to bacteria (or even human cells)! But how do you make recombinant plasmids in a laboratory? DNA is very, very long. How can you cut it into useful sized pieces, find the right pieces, stick them together, etc.?? The solution was discovered by pursuing a phenomenon known as restriction (described below), which seemed at the time to be of no practical consequence whatsoever.

B. Discovery of Restriction & Modification Enzymes See Becker Box 18B and/or Sadava 16.1

1. *The phenomenon:* Some phages grow well on certain bacteria but not others. For example:

a. Virus V infects bacteria of type A → lots of progeny.

b. Virus V infects bacteria of type B → very few progeny.

The growth of the phages (viruses) is said to be "restricted" on bacteria of type B (case b). But an occasional virus particle does manage to complete an infection and lyse a type B bacterium in spite of restriction. If it does, the progeny phage from the lysed cell grow just fine on type B -- they are no longer restricted. If you take one of the few progeny from (b) and infect bacteria of type B → lots of progeny.

What is surprising is that the sequence of bases in the viral DNA is the same before and after restriction! In other words, all the progeny of Virus V have the same DNA

sequence, whether you examine the original virus, the progeny from a or the progeny from b. What's going on here? The solution holds the key to all genetic engineering and recombinant DNA technology, and the scientist who discovered it (Arber) received the Nobel Prize in 1978 along with two other scientists (Smith & Nathans) who extended Arber's work.

2. Restriction enzymes. What causes restriction?

- What are restriction enzymes? Bacteria have endonucleases that cut DNA molecules at specific sequences. (All previously known endonucleases cut at random.) These enzymes are called restriction enzymes or restriction endonucleases.

  *Exo*nucleases remove nucleotides one at a time from the end.
  *Endo*nucleases break specific phosphodiester bonds in the middle of the molecule.

- What's the substrate? The enzymes work only on double stranded DNA. They don't cut RNA or single stranded DNA.
- Where do they cut? Each restriction enzyme catalyzes hydrolysis of a phosphodiester bond between two nucleotides. (Both strands of the DNA are cut at equivalent points.)
- Specificity: Each different restriction enzyme recognizes (binds to) and cuts the DNA at a different particular sequence. (See handout 17A-1 for some examples.)
- What is the normal function of these enzymes? To destroy the DNA of infectious phage.

3. Modification enzymes. Why are bacteria "immune" to their own restriction enzymes? Why don't the restriction enzymes cut up the bacterium's own DNA?

- Modification enzymes: Bacteria have a second set of enzymes that modify DNA -- these enzymes add methyl groups to specific sequences on the DNA -- the same sequences cut by the restriction enzymes.
- Role of modifications: These modifications make the bacterial DNA resistant to the restriction enzymes.
- What's the substrate? The modification enzymes normally add methyl groups only to DNA that is hemi-methylated -- methylated on one strand, but not the other.
- Inheritance: The state of methylation is heritable. When methylated DNA is replicated, the product is hemi-methylated -- the new strand has no methyl groups. The new strand is methylated by modification enzymes soon after it is made. When un-methylated DNA is replicated, no methyl groups are added. So the state of DNA (methylated or not) is maintained generation after generation, once it is set up.
- Terminology -- The inheritance of restriction (or the pattern of methylation) is said to be 'epigenetic.' The DNA sequence isn't changed, but the DNA is modified. The modified state of the DNA is stable, and the state is inherited.

4. How do rare virus particles escape restriction? If DNA is accidentally methylated, then it will be methylated every time it is replicated thereafter. The rare particles have DNA that was methylated accidentally. Therefore the DNA of the rare virus particles and their progeny are resistant to the restriction enzymes.

5. Modification in general: Modification of macromolecules by enzymatically adding or removing a small group or two is very wide spread, especially in eukaryotes. It is a common method of regulating and/or fine tuning macromolecular function permanently or temporarily (many modifications are reversible). For example, many enzymes are activated or inhibited by addition of phosphate groups; some sections of DNA may be kept "off" by addition of methyl groups to the DNA itself or to associated proteins (histones). Details of these examples (and many others) will follow next semester.

C. Examples & Properties of Restriction Enzymes (Details on Handout 17A-1 & Becker Box 18B)

1. Restriction Sites are often palindromes (= read the same forwards and backwards)

English examples:	"Madam I'm Adam"
	"Able was I ere I saw Elba."
DNA example:	5' GAATTC 3' 3' CTTAAG 5'

What do we mean by "reads the same backwards and forwards" for the DNA example? There are several ways to explain this :

a. By base sequences. The sequences on the two individual strands are the same if both sequences are read 5' to 3'.

b. By base pairs. The base pairs are the same, right to left and left to right as long as the "top" strand is always 5' to 3'.

c. By rotation. If you rotate the DNA 180 degrees, it looks exactly the same.

2. State of the ends. Cuts made by restriction enzymes can be staggered (generating so called "sticky ends") or blunt (see handout 17A or Becker Box 18B for examples)

3. Sites can sometimes be methylated -- this makes the sites resistant to cutting. (See Modification enzymes, above.)

4. There are a wide variety of restriction enzymes made by different bacteria. (See handout or texts for some examples.)
Therefore there are many different options for cutting up any given DNA. For example:

- Some enzymes recognize relative short sequences. For example, an enzyme may be a "4 cutter" = enzyme that recognizes a 4 base pair site. (See handout.) Short sites (sequences) are found more often, and enzymes that cut them produce many relatively short fragments.
- Some enzymes recognize longer sequences. Longer sites are found less often, and enzymes that cut them produce a smaller number of relatively long fragments.

D. Significance of Restriction enzymes = Essential tools for Recombinant DNA Analysis

1. Allows you to cut up DNA into manageable size pieces for manipulation and analysis.

a. Most DNA is very long. Without some sort of breakage, most DNA molecules are too big to handle.

b. Before this, all known DNases cut at random → big mess (random collection of different sized pieces).

c. Restriction enzymes cut DNA into fixed size pieces. Pieces resulting from restriction enzyme digestion can be separated by size using gel electrophoresis. (Sadava 16.2 or Becker fig. 18-12.) Principle is similar to SDS gel electrophoresis, except no SDS is used -- all nucleic acids are negatively charged and migrate to the positive pole -- smaller fragments travel farther. (Same procedure as used for analysis of PCR products.)

2. Joining. Existence of "sticky ends" allows you to join, not just cut, DNA's readily, using overlapping ends and ligase. This allows you to make new, recombinant molecules in a test tube. See handout 17A-3 (= Becker fig. 20-25 [20-26]) or Sadava 16.8 (16.4) Why is this helpful? Allows molecular cloning, as explained below.

3. Forensics/ IDs & RFLPs

a. Inherited variations in base sequence lead to differences in places where DNA is cut. For example, if a sequence is GAATTC, EcoR1 will cut the DNA. If the sequence is changed to GGATTC, EcoR1 will not cut the DNA. So a change of A to G can "remove" a restriction site while a change of G to A can "add" a restriction site. (See handout 17A-2.) Note "addition" or "loss" of a restriction site does not necessarily mean the insertion or deletion of bases. An "addition" means a change in base sequence so a particular stretch of DNA is now recognized by a particular restriction enzyme; a loss means a change so that the DNA is no longer recognized by the restriction enzyme.

b. Detection. Differences in restriction sites (or variation in the lengths of the sequences between sites) are detected by cutting the DNA with restriction enzymes and running gels of pieces (to compare sizes). How to locate the pieces from a particular region of the DNA will be discussed next time.

c. RFLPs. When DNA from different individuals gives a different pattern of pieces, this is known as a RFLP = restriction fragment length polymorphism. (See handout 17A-2 or Becker box 18C, or Sadava 17.8 for examples.)

d. How does this compare to PCR fingerprinting?

- PCR method for DNA fingerprinting (described previously) picks up differences in the numbers of repeats in a particular region.
- To use PCR, you have to know enough about the DNA to make the correct primers. No primers are needed for this method.
- This method picks up differences in length -- for any reason -- between restriction sites. Variations in length can be due to differences in the restriction sites themselves (as in **problem 13-3**) or to differences in the length of the sequences between the restriction sites.
- Additional uses of RFLPs will be explained in detail later, after human genetics.

To review restriction enzymes, try Problems 13-1 & 13-2 parts B & C, 13-3, & 13-8 A. (Problem 13-8 is not in older editions of the problem book; if you have an older edition, see the update page.) To review RFLPs try 13R-4, parts A & B. (Homozygous means both copies of the gene in a diploid, such as a human, are the same.)

II. Probes

A. What's a probe? It's a nucleic acid that's complementary to the sequence you are looking for, and it is usually labeled or tagged in some way -- with radioactivity, fluorescence, or something else that's relatively easy to detect. Probe may consist of single stranded DNA or double stranded DNA. (Double stranded DNA must be denatured before it will hybridize to the target DNA.)

B. Why would you need a probe?

To find the right piece of DNA -- the piece from a particular part of the genome. If you cup up genomic DNA, you will have many pieces of DNA, and you want to find all the ones carrying a particular sequence. For example, to detect an RFLP (such as the one shown on handout 17A-2), you need to look at DNA from that particular part of the genome. But if you cut up the total DNA of an organism, you get many, many pieces. How will you find the right pieces, that is, the ones carrying a particular gene or section of the DNA where the RFLP is?

C. How do you get a probe that is complementary to a particular gene?

1. cDNA (works well for abundant proteins). Use reverse transcriptase to make DNA copy of mRNA. (Use additional enzymes if you want ds cDNA. See Becker fig. 20-30 [20-31].) DNA copy of mRNA is called "cDNA." (cDNA = complementary DNA or copy DNA.) cDNA is commonly made and used as a probe for a highly expressed genes -- ones that make a lot of protein and therefore a lot of mRNA.

Question 1: Why use cDNA instead of mRNA itself? DNA is more resistant to degradation and easier to work with than RNA. See Sadava 16.12 (16.8).

Question 2: If you have the cDNA, why do you need the gene itself? The gene or genomic DNA is usually **not** the same as the cDNA, because of introns, as explained below.

2. Oligonucleotide probes (good if know amino acid sequence) -- Sequences up to 50-100 nucleotides long can be synthesized by chemical means. An oligonucleotide of 15-20 nucleotides is usually sufficient to act as a probe. (That is, the sequence hybridizes to the gene you want and not to other genes.) You actually need a mixture of oligonucleotides -- you need more than one because the code is degenerate. Can't predict exact DNA/mRNA sequence from amino acid sequence. For example, if amino acids are lys - asp - met etc, DNA would be AAG/A CAC/T ATG etc. Don't know if base #3 is G or A; # 6 is C or T etc. So you make a mixture of oligonucleotides -- some have G at position 3, and some have A etc. (During chemical synthesis use a 50:50 mixture of G and A at position #3.) You'll get a mix of oligonucleotides, and one of the combinations will be the right one to hybridize to your gene.

3. Brute force/chance -- cut up DNA and get random collection of pieces; then use some procedure to figure out which is right piece. This sounds insane, but has been used very successfully to locate some human disease genes, as will be explained later after we cover the necessary genetics.

IV. Blots

A. What's a Blot? -- using probes to detect DNAs that are immobilized on a solid support. DNA does not need to be purified first. Can be released from colonies in place (*in situ*) or "blotted" from a gel. Once DNA is stuck to a support, it can be denatured while still attached to the support. Then you can add a solution of probe (the complementary, labeled, DNA) and see if probe hybridizes to the denatured DNA. You wash off unattached (unhybridized) probe and see what is left. That allows you to identify band, colony, etc. containing the nucleic acid of interest (= nucleic acid that hybridizes to and/or traps probe.) Blots allow you to test hybridization of probe to many DNA samples at once. (How you use blots to test DNA from colonies will be discussed next time.)

D. Basic procedure for detecting DNA Bands on gels Handout 18A, right panel. Italics = terms on HO. How do you find a particular fragment of DNA? The one containing the gene or RFLP you are looking for? See Becker, box 18C or Sadava 16.3.

1. Cut DNA up, or do PCR to amplify selected pieces. (If you cut the DNA with restriction enzymes, either cut up the same DNA with several different enzymes, or cut up several different samples with the same enzyme.)

2. Separate pieces on gel (*agarose gel electrophoresis*). Consider: Without probe, what would pattern of bands look like? How can you find the band you want without cutting up the whole gel into slices and testing each one? A "Southern Blot" allows you to do this.

3. Blot DNA from gel to paper or plastic (*transfer to nitrocellulose*.) DNA sticks to the support.

4. Denature DNA (in situ)

5. Add probe (*labeled cDNA or RNA)*

6. Allow DNA and probe to hybridize in situ. Note: You can use stringent or nonstringent conditions, depending on whether you want a only a perfect match or are looking for approximate matches too.

7. Put blot in dark next to film if label is radioactive

8. Detect bands (*look at autoradiograph = developed film*; use other methods if label is non-radioactive.)

E. Types of Blots (from gels) -- Terminology & Variations -- Southerns, Northerns and Westerns.

1. Southerns -- Cut up DNA, separate DNA fragments by gel electrophoresis, blot, find desired fragment by hybridization to probe.

2. Northerns --Separate RNAs by gel electrophoresis. (RNAs are small enough as is; don't cut them up first.) Then blot, hybridize as before.

3. Westerns -- Separate proteins by SDS gel electrophoresis, blot, find desired protein using antibodies specific for that protein.

F. Example of use of Blots -- how to detect the RFLP's on handout 17A-2. Example will illustrate the following features:

1. Need not purify the DNA of interest first -- you test a mix and locate where in the mix your DNA of interest is. You separate everything first by size (on gels), and then find the position of the piece you are interested in/want. Or you grow up multiple colonies containing different DNA sequences. You don't have to know in advance where the sequence you want will be. Needle in the haystack, but it works! (You spread out the hay in your haystack and glue to a support. Then cover it with magnetic particles and shake off the ones that don't stick. Where there are particles, your needle is underneath. This analogy is nice but misses point that your "haystack" or "needles" are sorted by size when using gels.)

2. Can test many samples at once -- can use multiple wells and/or repeat hybridization to same blot using diff. probes.

3. Probe need not be same length as fragment -- probe can be shorter or longer than target sequence or fragment you are looking for. Probe and target need not be the same length, but there must be overlap, so some region of target hybridizes to probe. Sample must "capture" probe. (Think Velcro.)

To review blots & probes, try problem 13-8B. Do 13-8A first if you haven't done it already. You may want to wait to do 13-8B until after topic V.

V. Molecular (DNA) Cloning -- how do you isolate a gene and why bother?

A. What is DNA cloning? Means making a chimeric (recombinant) plasmid or virus and growing it up. For an overview of how to make a chimeric plasmid, see Becker 20-27 [20-28] or handout 17A (upper right). For an overview of the whole process (to be discussed next time) see Becker 20-26 (20-27).

B. Why bother?

1. To get a lot of copies of a particular gene/DNA.

a. To look at gene structure. Allows you to examine structure of gene and/or its regulatory sequences. There are some big surprises from this, as will be discussed next time.

b. Why is cloning necessary nowadays if you have PCR? Can't do PCR if don't know flanking sequences & can't make primers -- have to clone the first time.

2. To get a lot of gene product (protein, not DNA). Set up conditions where gene will be transcribed and translated. This allows you to

a. Measure gene expression (how much product a gene makes) and its regulation. Some examples of this next term.

b. Produce useful amounts of a product that is otherwise unobtainable in large quantities. For example, protein hormones such as growth hormone or erythropoietin (EPO). See Sadava 16.17 (16.14) for info on a similar example (TPA).

3. For gene therapy -- to restore function to a defective cell. Want to add gene so it will remain and supply a missing product as needed.

C. Why you need a " cloning vehicle" or "vector." (See "fragment vs. plasmid" on handout 16B.) Need to add your gene of interest to a plasmid or modified virus so your gene can be replicated (using the origin of replication of the plasmid or virus). If you use a virus, it is usually modified, so that the DNA you are cloning replaces some of the viral genes.

D. Basics of how to make a recombinant plasmid. You put the 'vector' and the fragment you wish to clone together using sticky ends (See handout 17A, top right.)

E. Details to know about plasmid/virus = cloning vehicle. See handout 17B (left half) or Becker fig 20-27a [20-28a] or Sadava 16.9 (16.5) -- What properties should the vector have?

1. Why you want 1 recognition site for each restriction enzyme per plasmid. You want 1 insertion site on each plasmid (for each dif. enzyme). There may be many different sequences recognized by dif. restriction enzymes, but only one site for each.

2. Why sites for many dif. enzymes -- so you can cut up DNAs in different ways and insert any of the fragments in the same plasmid.

3. Why drug resistant -- need a selectable "marker" on the plasmid -- gene that confers growth only to cells that got the plasmid. Then cells without the added "selectable marker" (that is, without a plasmid) won't grow under some condition. For example, the "selectable marker" can be a gene that confers resistance to the antibiotic tetracycline. (Tet-R) This allows you to select **for** cells that got a plasmid and **against** cells that got no plasmid at all. Cells without plasmids will be sensitive to the drug and won't grow in the presence of antibiotic. Only cells that have received a plasmid with the Tet-R gene will be drug resistant and will grow in the presence of antibiotic.

To review blots & probes, and genetic engineering so far, try problems 13-4 (A-D) & 13-8B if you haven't done it already. (After we explain introns and exons, try problem 13-11, esp. parts D & E.) Most of the "blot & probe" problems involve introns or other complications we haven't addressed yet.

4. Why usually double drug resistant or has second selectable marker (See Sadava 16.10 (16.6) or handout 17B) -- so you can detect cells that received a plasmid **without** an insert. The selectable marker (= active gene) confers some property that is easily detectable, for example, growth in the presence of an antibiotic or formation of blue colonies. If there is an insert in the 'selectable marker', that gene will be inactivated, and the cells will not have the corresponding property --

they will not grow, or not turn blue, etc. Conversely, if there is no DNA fragment inserted in the plasmid, the selectable marker gene will continue to function, and the cells will have the appropriate property -- they will continue to grow (plus antibiotic) or turn blue, etc. Examples of selectable markers:

a. Drug resistance. In the plasmid shown on the handout (pBR322), the second selectable marker is the gene for ampicillin (Amp) resistance. If the plasmid has no insert, it confers Amp resistance. But if the plasmid has an insert, Amp resistance is lost. You want to identify the cells that got a plasmid, but are NOT Amp resistant.

b. LacZ. Another common selectable marker is the lacZ gene. Normal cells take up and cleave an analog of lactose; the product turns the colonies blue. Cells with an inactivated lacZ gene do not split the analog and remain colorless. (See Becker fig. 20-27 [20-28] & accompanying text).

Next Time: Wrap up of genetic engineering -- details of how you make a recombinant plasmid, and how you find a cell that got the gene you are trying to clone. Then on to Eukaryotic Cell Division. How does one eukaryotic cell make two?