7.016 Recitation 11 – Fall 2018

*(**Note:** The recitation summary should NOT be regarded as the substitute for lectures)*

Summary of Lecture 16 (10/17):

Recombinant DNA & Cloning: Recombinant DNA contains fragments from two different sources of DNA, either from the same or from different species. The common reagents necessary for recombinant DNA cloning are restriction enzymes, DNA ligases and vectors.
https://www.youtube.com/watch?v=acKWdNj936o&feature=related

Restriction enzymes: These endonucleases occur naturally in bacteria and protect them against viral infections. These enzymes have been co-opted by scientists for use in molecular biology. Restriction enzymes cut the DNA by recognizing and binding to specific restriction sites. Most of the recognition sites for the restriction enzymes are 4 to 6 base pairs (bp) long and are palindromes meaning that both strands of the DNA have the same base sequence when they are read in the 5'→3' direction. Restriction enzymes can either be staggered cutters (which generate 5' or 3' overhangs) or blunt cutters (which do not generate any overhang). While selecting the restriction enzymes for cloning, scientists often use a set of enzymes that allows directional cloning of the insert into the plasmid.

DNA ligase: This enzyme is used to join the DNA fragments that get replicated on the lagging strand. The ligases can also be used to join the pieces of DNA, which have been cut by the same restriction enzymes or by the set of restriction enzymes that generate complementary ends.

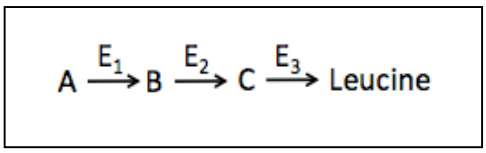
Vectors are the pieces of DNA that possess the qualities that allow any piece of DNA to be taken up by the cell and get replicated using the cell replication machinery. They can either be linear (viruses) or circular (bacterial or yeast plasmids). The most common vector is a **plasmid**, which is a small circular double stranded DNA. Plasmids used as vectors to carry DNA into bacterial cells should have the following features:

- Recognition sequence for a restriction enzyme(s) so that you can cut the plasmid open with these enzymes and insert a piece of DNA (which can either be a gene or a cDNA (often called gene insert or cDNA insert)) into it.
- An origin of replication so that the plasmid can replicate after it is within the bacterial cell
- A marker gene such as an antibiotic resistant gene (Amp^r, Tet^r, Kan^r etc) so that you can select for the bacterial cells that took the plasmid or a reporter gene such as lacZ (which encodes an enzyme which converts a white color substrate to blue) or a green fluorescence protein (GFP) gene so that you can screen for the bacterial cells that are expressing the gene.

Transformation, selection and screening: Once you have the recombinant plasmid with the DNA sequence of interest (called the insert), you transform the plasmid into the competent bacterial cells. You then grow the transformed bacteria on a plate that contains media that allows the growth of both transformed and untransformed cells (master plate). Then you replica plate the cells on plates that contain minimal media, which allows you to select for the bacterial cells that took the plasmid with the insert from those that did not. For example, you can grow the transformed bacterial cells on a plate that contains specific antibiotic to which the gene on the plasmid confers resistance. Any cell that took up the plasmid will grow on medium containing this antibiotic and those that did not take in the plasmid will die in the medium, thus allowing selection of the transformed colonies. If you identify the transformed colonies by looking at the expression of a reporter gene (such as *lacZ* which encodes β-galacosidase, an enzyme that transforms white X-gal substrate to blue color precipitate) then you say that you have screened for colonies that have the insert.

DNA gels: DNA gels are slabs of materials such as agarose that form gelatinous matrices when polymerized. DNA or RNA can be inserted into such gels, the gels can be immersed in liquid, and then an electrical current can be applied to the gel such that these macromolecules move towards the pole to which they are attracted. Very long macromolecules will move slowly as they attempt to weave their way through the pores in the gel. Smaller macromolecules will move faster. This technique of gel electrophoresis allows macromolecules to be separated by size. Since DNA is negatively charged it always moves from negative electrode to the positive electrode.

Cloning by complementation (Functional cloning): Here you clone a gene that can complement the mutation within a cell or an organism. For example, let us look at the biochemical pathway.



In the above biochemical pathway, the enzyme E_1 (encoded by Gene 1) catalyzes the conversion of A→B, E_2 (encoded by Gene 2) catalyzes the conversion of B→C, and E3 (encoded by Gene 3) catalyzes the conversion of C→Leucine. An $E1^{+}E2^{+}E3^{mut}$ cannot synthesize Leucine (auxotroph) and can therefore not grow in a growth medium that lacks Leucine. If this mutant is transformed with a recombinant plasmid that has the wild-type copy of Gene 3, then the mutant will be able to synthesize Leucine (prototroph) and grow in a medium that lacks Leucine.

Making genomic Library: A library is a collection of different recombinant DNA molecules (often stored in bacterial or yeast cells), the set of which represents all of the genetic material of an organism. A mouse genomic library, for example, would be a population of host bacteria, each of which carries a piece of mouse genomic DNA that was inserted into a cloning vector, such that the collection of cloned DNA molecules represents the entire genome of the mouse in bacteria.

Making cDNA Library: An alternative to a genomic library is a cDNA library. A cDNA library represents, not the entire genome, but only the DNA that makes the reading frames of the expressed genes. cDNA is complementary DNA, which is DNA that was made in the laboratory by isolating total mature mRNA from the host organism and reverse transcribing each mature mRNA molecule into a double-stranded DNA molecule using a PolyT primer that can base pair with the poly A tail at the 3' end of the mature mRNA. Each cDNA is then cloned into an appropriate vector, the bacterial cells are then transformed with these recombinant vectors and the set of recombinant molecules in bacteria is referred to as the mouse cDNA library in the bacteria.

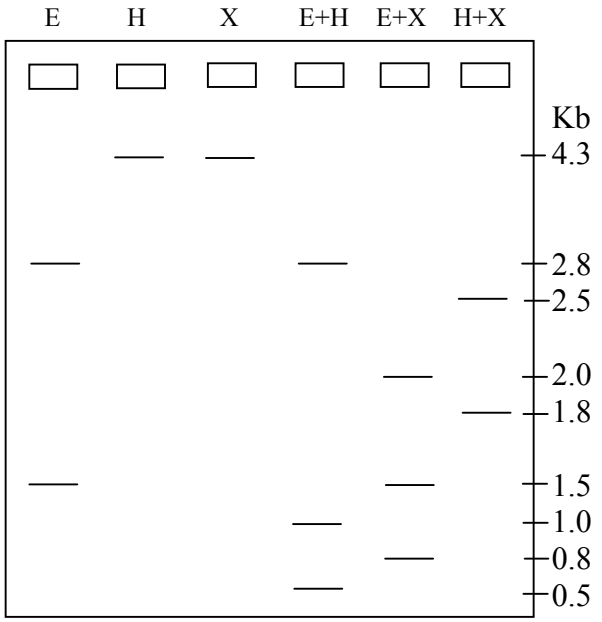
It is important to note that cDNA copy of a gene does NOT contain the promoter. Therefore if you clone a cDNA form of the gene in a plasmid and would like to express it in bacterial cells, your plasmid should have a bacterial promoter.

https://www.youtube.com/watch?v=SvjeCxVu2dI

Fusion gene: It is a hybrid gene that is formed from two separate genes. For example you can fuse the Green fluorescent protein gene (GFP gene) to the start (sequence corresponding to the N terminus) or the end (sequence corresponding to the C terminus) of any gene of interest (GOI) to create a GFP-GOI or GOI-GFP fusion gene. This fusion should maintain the reading frames of both genes and there should not be any stop codon in between. This encodes of a fusion protein (GFP-GOI fusion protein or GOI-GFP fusion protein). Creating such proteins allow us to study the function of a particular protein within a live cell or an organism under different conditions.

Questions:

1. Three restriction enzymes have recognition sites in a plasmid: EcoRI ("E"), HindIII ("H"), and XbaI ("X"). You digest the plasmid with each of the following combinations of enzymes and see the following gel.



a) Draw a map of the plasmid indicating where each restriction enzyme cut site is, which restriction enzyme cuts at each site, and how far apart each cut site is. Also give the total size (in kb) of the plasmid.

b) What basic features should this plasmid have to serve as a vector for cloning and expressing a cDNA copy of a human gene in bacteria?

2. You plan to make a PKX-GFP fusion gene. You start by fusing the cDNA corresponding to the N-terminus of GFP protein to the cDNA encoding the C terminus of PKX gene.

The following is the partial **cDNA sequence** encoding the C terminus of the PKX gene. The sequence encoding the stop codon is shown in bold and underlined. The bars above the sequence show the recognition sites for restriction enzymes Z and X.

```
                  Z                       Y
5'-TC AAG AGG ATC CCC GCG GTA CCG AAT TCC ATG TTA TAG CAA-3'
3'-AG TTC TCC TAG GGG CGC CAT GGC TTA AGG TAC AAT ATC GTT-5'
```

The following is the partial cDNA sequence encoding the N terminus of GFP. The codons are underlined. The bars above the sequence show the restriction enzyme recognition sites.

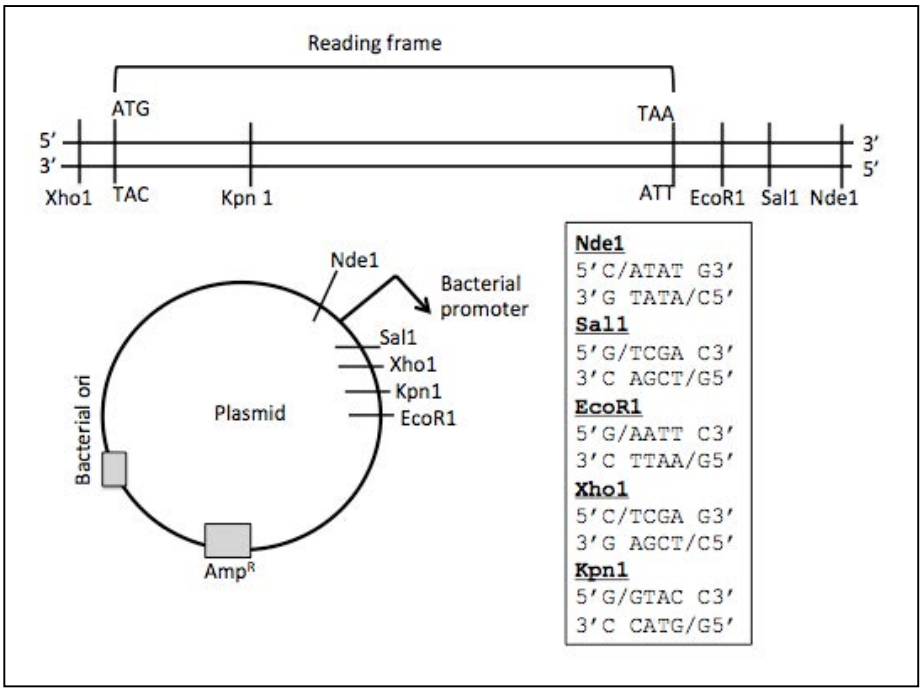
```
                          Z       Y
5'-TCT AGA GGT ACC GGG ATC CGA ATT CCC GTG CCA AGC GGC-3'
3'-AGA TCT CCA TGG CCC TAG GCT TAA GGG CAC GGT TCG CCG-5'
```

The recognition sequences and the cleavage sites (indicated by /) for each enzyme are given below.

Enzyme Z	*Enzyme Y*
5' G/GATC C 3'	5' G/AATT C 3'
3' C CTAG/G 5'	3' C TTAA/G 5'

Choose the restriction enzyme that you will use to cut the two genes before ligating them to make a fusion gene. **Explain** why you chose this restriction enzyme.

3. You want to clone and express a cDNA corresponding to human gene in bacterial cells using a plasmid diagrammed below. Below is a restriction map of a portion of the sequence that contains the cDNA of your interested.



Your task is to design a strategy to insert the yeast gene into the bacterial plasmid. For each pair of restriction enzyme, explain whether you can use them to clone and express the yeast genomic DNA and the plasmid.

a) Pair 1: NdeI & XhoI

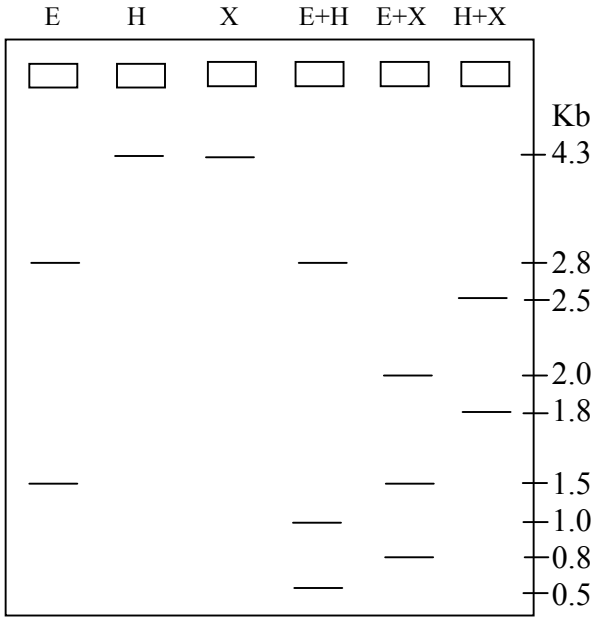
b) Pair 2: SalI & KpnI

c) Pair 3: SalI & XhoI

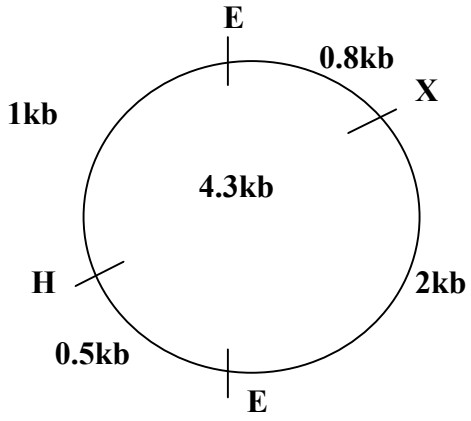
d) Pair 4: XhoI & EcoRI

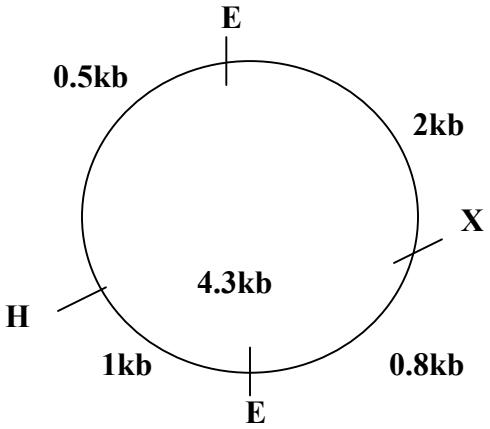
The key:

1.Three restriction enzymes have recognition sites in a plasmid: EcoRI ("E"), HindIII ("H"), and XbaI ("X"). You digest the plasmid with each of the following combinations of enzymes and see the following gel.



a) Draw a map of the plasmid indicating where each restriction enzyme cut site is, which restriction enzyme cuts at each site, and how far apart each cut site is. Also give the total size (in kb) of the plasmid.





b) What basic features should this plasmid have to serve as a vector for cloning and expressing a cDNA copy of a human gene in bacteria?

*To clone a gene in a plasmid, the plasmid should have an origin of replication, a restriction enzyme site so that the plasmid can be cut open and used as a vector to clone the desired sequence and a marker gene (i.e. antibiotic resistant gene) that can be used to differentiate between the untransformed host cells and the host cells that have obtained the plasmid. (**Note:** If you are cloning an intact gene, then it has its own inherent promoter. So it can be expressed in specific cell i.e. yeast gene can be expressed in yeast but not in bacterial cell. So your plasmid does not need a promoter sequence. However, if you also want to express the cloned cDNA you will need a species specific promoter prior to the transcription start site since the cDNA will lack the promoter sequence).*

2. You plan to make a PKX-GFP fusion gene. You start by fusing the cDNA corresponding to the N-terminus of GFP protein to the cDNA encoding the C terminus of PKX gene.

The following is the partial **cDNA sequence** encoding the C terminus of the PKX gene. The sequence encoding the stop codon is shown in bold and underlined. The bars above the sequence show the recognition sites for restriction enzymes Z and X.

```
                  Z                          Y
5'-TC AAG AGG ATC CCC GCG GTA CCG AAT TCC ATG TTA TAG CAA-3'
3'-AG TTC TCC TAG GGG CGC CAT GGC TTA AGG TAC AAT ATC GTT-5'
```

The following is the partial cDNA sequence encoding the N terminus of GFP. The codons are underlined. The bars above the sequence show the restriction enzyme recognition sites.

```
                        Z       Y
5'-TCT AGA GGT ACC GGG ATC CGA ATT CCC GTG CCA AGC GGC-3'
3'-AGA TCT CCA TGG CCC TAG GCT TAA GGG CAC GGT TCG CCG-5'
```

The recognition sequences and the cleavage sites (indicated by /) for each enzyme are given below.

Enzyme Z	*Enzyme Y*
5' G/GATC C 3'	5' G/AATT C 3'
3' C CTAG/G 5'	3' C TTAA/G 5'

Choose the restriction enzyme that you will use to cut the two genes before ligating them to make a fusion gene. **Explain** why you chose this restriction enzyme.

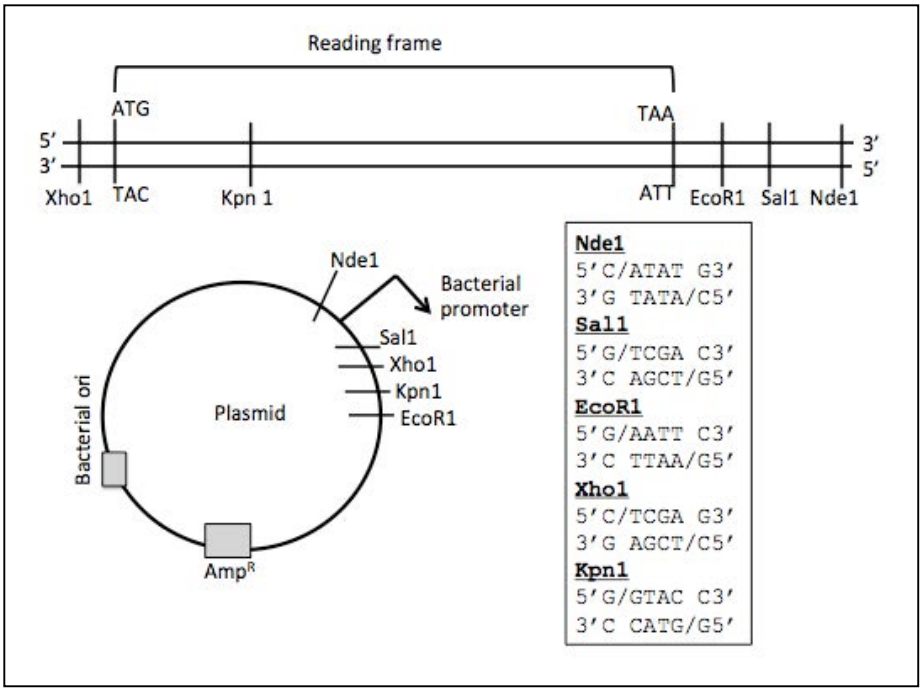
You CAN use restriction enzyme Z since it keeps the codons of the fusion gene in frame.

```
5'-TC AAG AGG ATC CGA ATT CCC GTG CCA AGC GGC-3'
3'-AG TTC TCC TAG GCT TAA GGG CAC GGT TCG CCG-5'
```

You cannot use the restriction enzyme Y since it puts the GFP cDNA out of frame.

```
5'-TC AAG AGG ATC CCC GCG GTA CCG AAT TC CCG TGC CAA GCG GC-3'
3'-AG TTC TCC TAG GGG CGC CAT GGC TTA AG GGC ACG GTT CGC CG-5'
```

3. You want to clone and express a cDNA corresponding to human gene in bacterial cells using a plasmid diagrammed below. Below is a restriction map of a portion of the sequence that contains the cDNA of your interested.



Your task is to design a strategy to insert the yeast gene into the bacterial plasmid. For each pair of restriction enzyme, explain whether you can use them to clone and express the yeast genomic DNA and the plasmid.

a) Pair 1: NdeI & XhoI: *Since you want to transcribe and express the cDNA in the bacterial cells, you will need a bacterial promoter (as shown on the plasmid). However if you use the restriction enzyme Nde1 and Xho 1 pair, you will delete this promoter. So this is not a good choice.*

b) Pair 2: SalI & KpnI: *You will not use this pair since kpn1 restriction enzyme cuts within/ disrupts the open reading frame of the yeast gene which you want to insert in the plasmid and express in the bacterial cells.*

c) Pair 3: SalI & XhoI: *Both Xho1 and Sal1 enzymes will generate sticky ends that are complementary to each other. So there is a 50% chance that the clone sequence will orient correctly with respect to the promoter and will be expressed there is also a 50% chance that it will orient opposite to the direction of the promoter and will not be expressed.*

d) Pair 4: XhoI & EcoRI: *This is the best choice since the cDNA will be inserted in correct orientation with respect to the promoter and will always be expressed.*

7.016 Introductory Biology
Fall 2018