7.013 Recitation 8 – Spring 2018

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

Summary of Lectures 13 (3/9) and 14 (3/12):

Recombinant DNA & Cloning: This is a DNA that 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 a specific restriction sites. Most of the recognition sites for the restriction enzymes are 4 to 6 base pairs (bp) long. Because each sequence of bases has a unique structure, it can be specifically recognized by a particular restriction enzyme. Most recognition sites for restriction enzymes are palindromes meaning that both strands of the DNA have the same base sequence when they are read from the 5' (or the 3') ends. 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.
https://enzymefinder.neb.com/#!#nebheader

DNA ligase: This enzyme is used to join the DNA fragments that get replicated on the lagging strand. It does so by forming the phosphodiester bonds between 3-OH'->5PO_4^{3-} (3'->5' phosphodiester bond)' between the replicated fragments of DNA. 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 cell should have the following features: a recognition sequence for a restriction enzyme(s) such 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 and a reporter gene such as an antibiotic resistant gene (Amp^r, Tet^r, Kan^r etc) that allow you select for the bacterial cells that took the plasmid or beta galactosidase gene (β- gal) or green fluorescence protein gene (GFP) that allows you to 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 bacterial cell (i.e. change the growth conditions to encourage the bacterial cells to take up recombinant plasmid). 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 that allows you to select 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 however you identify the transformed colonies by looking at the expression of the reporter gene (such as *lacZ* that encodes β-galacosidase enzyme which transforms white X-gal substrate to blue color precipitate) you say that you have screened for colonies that have the insert.

DNA gels: Gels are slabs of materials such as agarose and acrylamide that form gelatinous matrices when polymerized. DNA, RNA, and proteins 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. It is worth noting that 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 say you are looking at the biochemical pathway responsible for the synthesis of amino acid leucine: A-> B->C-> Leucine where the conversion of A->B is catalyzed by enzyme E1, B->C by enzyme E2, C->Leucine by enzyme E3, where E1, E2 and E3 are encoded by Genes *e1, e2 and e3*. An $E1^{+}E2^{+}E3^{mut}$ cannot synthesize leucine and can therefore not grow in a growth medium that lacks leucine (auxotroph). If this mutant is transformed with a recombinant plasmid that has the wild-type copy of the E3 gene insert, 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 would be 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 copying 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.

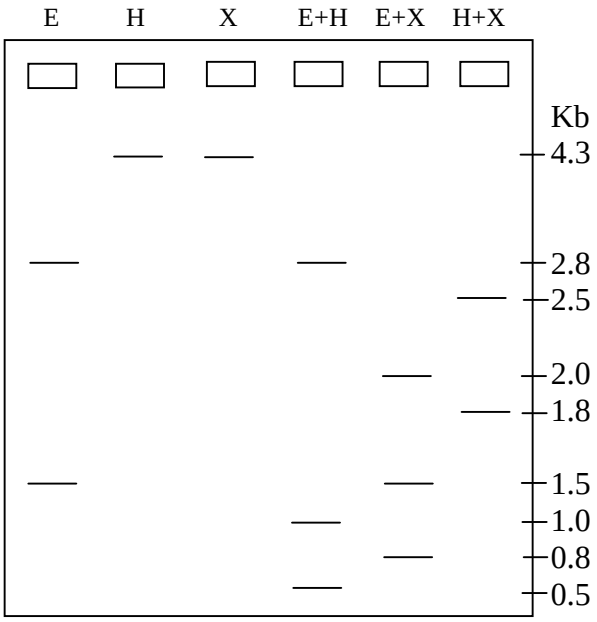
It is important to note that cDNA copy of a gene does NOT contain the promoter. Therefore if you clone a cDNA form of the gene in a plasmid and would like to express it in bacterial cells, your plasmid should have a bacterial promoter.

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 the gene of your 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 be any stop codon in between. This would encode 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.

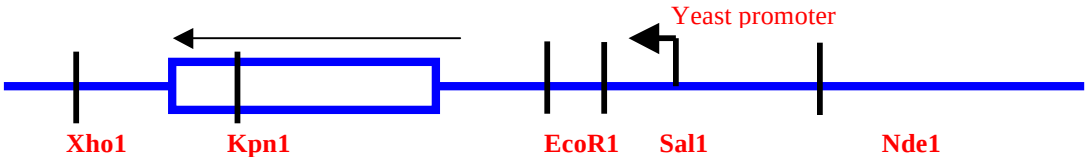


a) Draw a map of the plasmid indicating where each restriction enzyme cut site is, which restriction enzyme cuts at each site, and how far apart each cut site is.

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 want to insert a specific yeast gene (Gene A) into a specific bacterial plasmid such that the yeast gene will be transcribed in the bacterial cell. Below is a restriction map of a portion of yeast chromosome that contains Gene A in which you are interested. The box indicates the **open reading frame** of this gene.

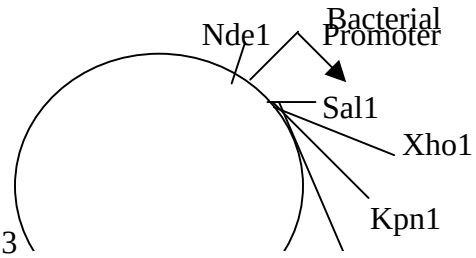
Gene is normally transcribed this way as shown by the arrow:



Below are the enzymes you can use, with their specific cut sites shown as "/"

***Nde* I:**	***Sal* I:**	***EcoR* I:**	***Xho* I:**	***Kpn* I:**
C/ATATG	G/TCGAC	G/AATTC	C/TCGAG	G/GTACC
GTATA/C	CAGCT/G	CTTAA/G	GAGCT/C	CCATG/G

On the right is the restriction map of the plasmid.



a) Your task is to design a strategy to insert the yeast gene into the bacterial plasmid. Which of the following pair of enzymes would you choose to cut the yeast genomic DNA and the plasmid.

i. NdeI & XhoI

ii. SalI & KpnI

iii. SalI & XhoI

iv. XhoI & EcoRI

b) If you did the digestion and ligation with the two enzymes you chose above, in how many ways could the insert be inserted into the vector? Draw the resulting recombinant plasmid with the yeast gene insert.

c) If the insert gets inserted backwards, what would the DNA sequences be at the two sites where ligation happened?

3. You plan to clone a PKX fusion gene. You start by fusing the DNA encoding GFP (green fluorescent protein) to the DNA 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. The bars above the sequence show restriction enzyme recognition sites.

```
          Z                    Y
5′-TCAAGAGGATCCCCGCGGTACCGAATTCCATGTTA**TAG**CAAGCTCGGAATTAACCCTCAC-3′
3′-AGTTCTCCTAGGGGCGCCATGGCTTAAGGTACAAT**ATC**GTTCGAGCCTTAATTGGGAGTG-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′-TCTAGAGGTACCGGGATCCGAATTCCC GTG CCA AGC GGC-3′
3′-AGATCTCCATGGCCCTAGGCTTAAGGG 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.

7.013 Introductory Biology

Spring 2018