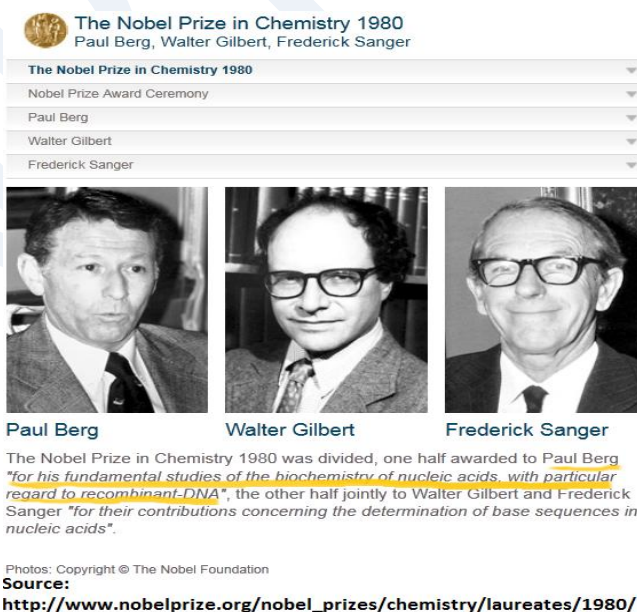# Lecture 35: Basics of DNA Cloning-I

Basics of DNA Cloning will be covered in two lectures during this course.

Cloning is making of identical copies. DNA cloning is process of making several identical copy of a gene or gene fragment. DNA fragment from an organism is cleaved or amplified and inserted in a DNA carrier called vector. Vectors are generally double stranded closed circular DNA which has origin of replication through which they can replicate in the host system. Vectors also have a selectable marker (generally antibiotics resistance gene) for screening of recombinant colonies. Vector with desired DNA insert is called recombinant DNA. This can be transferred to suitable host system (generally *E.Coli*) where it finds machinery for replication and makes several copies of it (may also express protein). The process is also called recombinant DNA technology or genetics engineering. Recombinant DNA technology is largely based on the work of Paul Berg, Herbert W. Boyer and Stanley N. Cohen although many other scientists have also made important contributions. Paul Berg in 1972, isolated a gene from a human cancer-causing monkey virus (SV40) using a restriction enzyme and joined this virus DNA with a molecule of DNA from the bacterial virus lambda using an enzyme called DNA ligase. This way the first recombinant DNA molecule was created. Later on, in 1980, Paul Berg shared Nobel prize in Chemistry for the work. A simplified concept of cloning is given in Fig. 1.
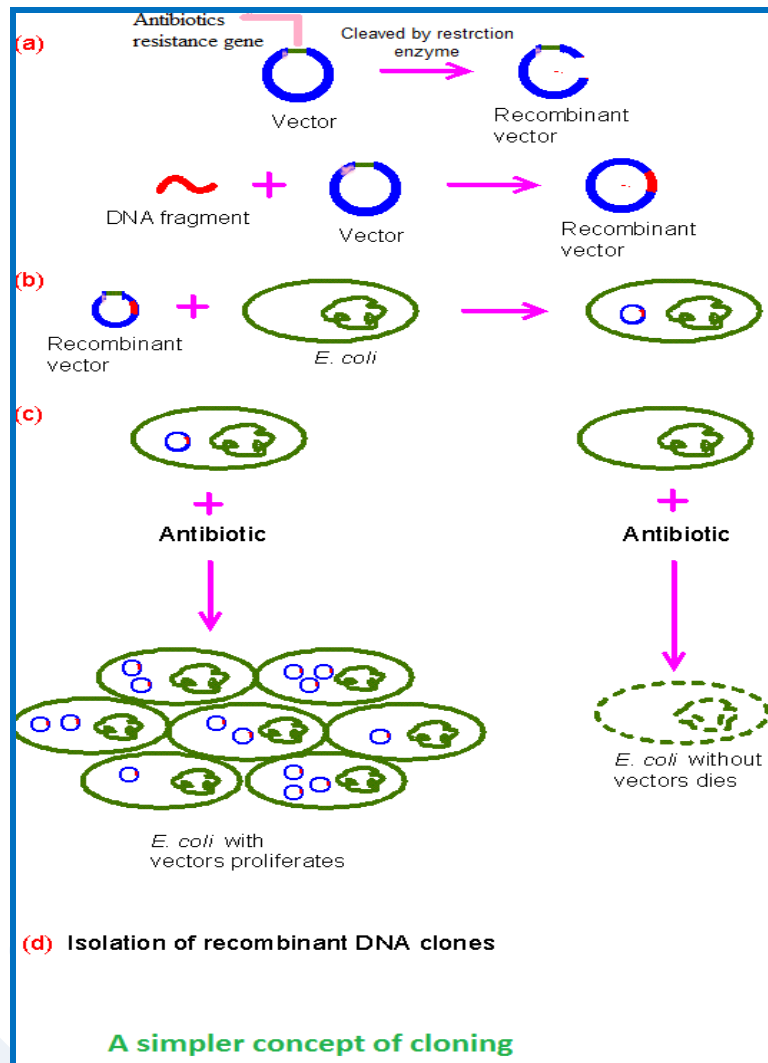

The Nobel Prize in Chemistry 1980
Paul Berg, Walter Gilbert, Frederick Sanger

The Nobel Prize in Chemistry 1980	▼
Nobel Prize Award Ceremony	▼
Paul Berg	▼
Walter Gilbert	▼
Frederick Sanger	▼

Paul Berg   **Walter Gilbert**   **Frederick Sanger**

The Nobel Prize in Chemistry 1980 was divided, one half awarded to Paul Berg *"for his fundamental studies of the biochemistry of nucleic acids, with particular regard to recombinant-DNA"*, the other half jointly to Walter Gilbert and Frederick Sanger *"for their contributions concerning the determination of base sequences in nucleic acids"*.

Photos: Copyright © The Nobel Foundation

Source:
http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1980/

Figure 1: A simplified concept of cloning

Cloning is a natural process in biology where genetically identical individuals are produced by asexually reproducing organisms such as bacteria, insects or plants. In biotechnology, the process of producing multiple identical copies of DNA fragments (molecular cloning), cells (cell cloning), or organisms is referred to as cloning. A clone has an exact genetic imprint as that of the original cell, tissue or organism. There are different types of cloning technologies used for various purposes besides producing the genetic copy of an organism. Basically the cloning technology can be divided into three types as reproductive cloning, therapeutic cloning and recombinant DNA technology or DNA cloning. Reproductive cloning is a technology used to

generate a twin of an animal that is genetically same as another currently or previously existing animal. The best example for reproductive cloning is Dolly, the first cloned sheep. Therapeutic cloning which is also known as "embryo cloning," is production of human embryos for use in research and treatment of diseases. The aim of this technique is not human cloning, but rather to harvest stem cells that are used for research studies and to treat diseases. The last and most widely used cloning technique in biotechnology is recombinant DNA technology.

In Biotechnology the gene is the cornerstone of most molecular biology studies. The study of genes can be facilitated by isolation and amplification of gene of interest. Cloning is one method used for isolation and amplification of gene of interest. The gene is cloned by inserting it into another DNA molecule which acts as vehicle or vector that will replicate in living cells. As the two DNA molecules of different origin are combined, the resulting DNA is known as recombinant DNA molecule. The term "gene cloning," "DNA cloning," "molecular cloning," and "recombinant DNA technology" all refer to same technique: Insertion of DNA fragment of interest from one organism into a vector which is a self- replicating genetic element inside a living cell. Gene cloning processes include removal of DNA from the cell, carrying out the DNA manipulations in test vial and, transformation of constructed DNA molecule back into the cells.

The first step in cloning is to prepare large amount of the vector and chromosomal DNAs. To carry the gene or the desired DNA fragment to the cell there is a need of a vector molecule. All cloning vectors are carrier DNA molecules. These carrier molecules host few common features in general such as; all vectors are self replicating in the cell, they contain a number of unique restriction enzyme cleaving sites that are present only once in the vector, they carry the selectable marker gene which is useful in selection of clone (usually an antibiotic resistance gene that is absent in the host cell) and, they can be very easily isolated from host cell. Depending on the purpose of cloning there are many vectors available. For use in the bacterial host *E. coli* system a greatest variety of cloning vectors have been developed. Thus, the first thing in cloning that a molecular biologist requires is to grow pure culture and isolate the cloning vector from the cells.

Choice of vector is dependent on insert size and application

The most commonly used cloning vectors include plasmids and bacteriophages (phage λ) beside all the other available vectors (Table 1). The cloning vectors are limited to the size of insert that they can carry. Depending on the size and the application of the insert the suitable vector is selected. The different types of vectors available for cloning are plasmids, bacteriophages, bacterial artificial chromosomes (BACs), yeast artificial chromosomes (YACs) and mammalian artificial chromosomes (MACs).

Plasmids: Plasmids are extra chromosomal circular double stranded DNA replicating elements present in bacterial cells. Plasmids show the size ranging from 5.0 kb to 400 kb. Plasmids are inserted into bacterial calls by a process called transformation. Plasmids can accommodate an insert size of upto 10 kb DNA fragment. Generally plasmid vectors carry a marker gene which is mostly a gene for antibiotic resistance; thereby making any cell that contains the plasmid will grow in presence of the selectable corresponding antibiotic supplied in the media.

Bacteriophage: The viruses that infect bacteria are called bacteriophage. These are intracellular obligate parasites that multiply inside bacterial cell by making use of some or all of the host enzymes. Bacteriophages have a very high significant mechanism for delivering its genome into bacterial cell. Hence it can be used as a cloning vector to deliver larger DNA segments. Most of the bacteriophage genome is non-essential and can be replaced with foreign DNA. Using bacteriophage as a vector, a DNA fragment of size up to 20 kb can be transformed.

Bacterial artificial chromosomes (BACs): Bacterial artificial chromosomes (BACs) are simple plasmid which is designed to clone very large DNA fragments ranging in size from 75 to 300 kb. BACs basically have marker like sights such as antibiotic resistance genes and a very stable origin of replication (ori) that promotes the distribution of plasmid after bacterial cell division and maintaining the plasmid copy number to one or two per cell. BACs are basically used in sequencing the genome of organisms in genome projects (example: BACs were used in human genome project). Several hundred thousand base pair DNA fragments can be cloned using BACs.

Yeast artificial chromosomes (YACs): YACs are yeast expression vectors. A very large DNA fragments whose sizes ranging from 100 kb to 3000 kb can be cloned using YACs. Mostly YACs are used for cloning very large DNA fragments and for the physical mapping of complex genomes. YACs have an advantage over BACs in expressing eukaryotic proteins that require post translational modifications. But, YACs are known to produce chimeric effects which make them less stable compared to BACs.

Human artificial chromosomes (HACs): Human artificial chromosomes (HACs) or mammalian artificial chromosomes (MACs) are still under development. HACs are microchromosomes that can act as a new chromosome in a population of human cells. HACs range in size from 6 to 10 Mb that carry new genes introduced by human researchers. HACs can be used as vectors in transfer of new genes, studying their expression and mammalian chromosomal function can also be elucidated using these microchrosomes in mammalian system. Different types of vectors are summarized in Table 1.

Table 1: Different types of vector with their properties.

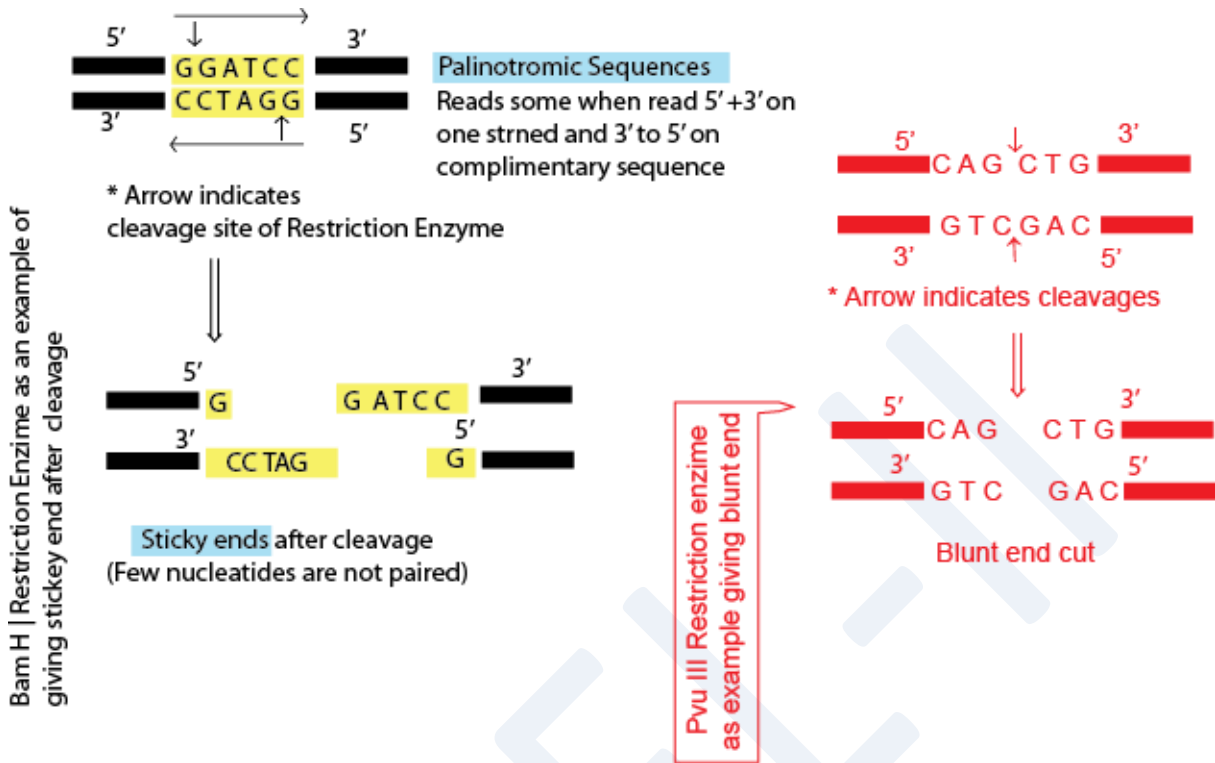| Vector	Basis	Size limit of insert	Major application
Plasmid	Naturally occurring multi copy plasmids	≤ 10 kb	Subcloning and gene manipulation, cDNA cloning and expression studies.
Phage	Bacteriophage λ	10- 20 kb	Genomic DNA cloning, cDNA and expression libraries.
Cosmid	Plasmid containing a bacteriophage λ *cos* site	35- 45 kb	Genomic library construction.
BACs	*Escherichia coli* F factor plasmid	75- 300 kb	Analysis of large genomes.
YACs	Saccharomyces cervisiae centromere, telomere, and autonomously replicating sequence	100- 3000 kb	Analysis of large genomes, YAC transgenic mice.
MACs	Mammalian centromere, telomere, and origin of replication	4> 10 mb	Still in budding stage for use in animal biotechnology and human gene therapy.

Cutting and joining DNA molecules:

To construct the plasmid with desire DNA fragment, both the plasmid and the desired DNA fragment has to be digested with the same restriction enzyme. Restriction enzymes are nucleases that cut double stranded DNA at specific nucleotide sequence known as restriction site. Theses enzymes were discovered in early 1960s by Werner Arber as DNA cutting enzymes. These enzymes are isolated by several types of bacterial species. It is interesting to note that a restriction enzyme found in a bacterial species does not cleave its own DNA, as recognition sites of the restriction enzymes are modified in the species by an enzyme called methylases. These restriction endonucleases in bacterial system are thought to have evolved as defense system to fight against foreign DNA that invading the bacterial cell such as viruses. To join the digested plasmid and insert, another enzyme called DNA ligase is require. Hence to generate recombinant DNA molecule two major categories of enzymes are required: restriction endonuclease and DNA ligase. Restriction enzyme cleaves phosphodiester bonds at the specific restriction sites and DNA ligases ligate double stranded DNA molecule by formation of phosphodiester bond between the two distinctly originated DNA molecules, thereby forming the recombinant DNA molecule which is transformed to bacterial cells. Different types of restriction endonucleases with their properties are summarized in Table 2.

Table 2: Classification of restriction endonucleases

Class	Abundance	Recognition site	composition	Use in recombinant DNA technology
Type I	Less common	Cut both strands at a nonspecific location > 1000 bp away from recognition site	Individual recognition, endonuclease, and methylase activity.	Not useful
Type II	More common	Cuts double strand at a specific, usually palindromic, recognition site 4-8 bp	Endonuclease and methylase are separate, single-subunit enzymes.	Very useful
Type III	Rare	Cleaves single strand, 24- 26 bp downstream of the 3' recognition site.	Endonuclease and methylase are separate.	Not useful

Recognition sequences for type II restriction endonucleases:

Type II restriction endonucleases are homodimeric polypeptide. These homodimer enzymes recognize short nucleotide sequences of about 4-8 bp known as restriction site and are usually palindromic in nature (Fig. 2). Most of the restriction enzymes used in molecular biology research are six base cutters. Restriction enzymes such as *Eco*R1, cuts the double stranded DNA at its recognition site in which the single stranded complementary tails called "sticky" or cohesive ends are generated. These single stranded sticky ends can form hydrogen bond with the complementary DNA sequence from different source. For example, two DNA sequences of different origin both containing *Eco*R1 restriction site can be ligated if they are digested with the *Eco*R1 restriction enzyme, as both produce sticky ends that are complementary to each other (Fig. 3). Some of the type II restriction endonucleases, such as *sma*I, cuts double strand DNA at the same position and generate blunt ends when they cleave the DNA. Restriction endonucleases show high degree of specificity in recognizing specific restriction site which is specific for that particular restriction enzyme. Any change in recognition site of the restriction endonuclease essentially eliminates total enzymatic activity.

Palinotromic Sequences
Reads some when read 5'+3' on one strned and 3' to 5' on complimentary sequence

* Arrow indicates cleavage site of Restriction Enzyme

5' G G A T C C 3'
3' C C T A G G 5'

Bam H I Restriction Enzime as an example of giving stickey end after cleavage

5' G   G A T C C 3'
3' C C T A G   G 5'

Sticky ends after cleavage
(Few nucleatides are not paired)

* Arrow indicates cleavages

5' C A G C T G 3'
3' G T C G A C 5'

Pvu III Restriction enzime as example giving blunt end

5' C A G   C T G 3'
3' G T C   G A C 5'

Blunt end cut

Few other Examples at Restriction Enzyme

5' G A A T C 3'
3' C T T A A G 5'
Eco RI

5' G T A T C 3'
3' C T A T A G 5'
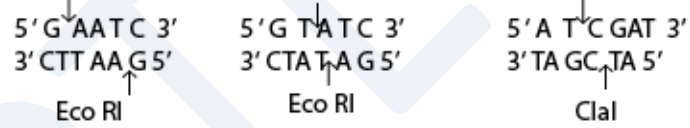Eco RI

5' A T C GAT 3'
3' TA GC TA 5'
ClaI

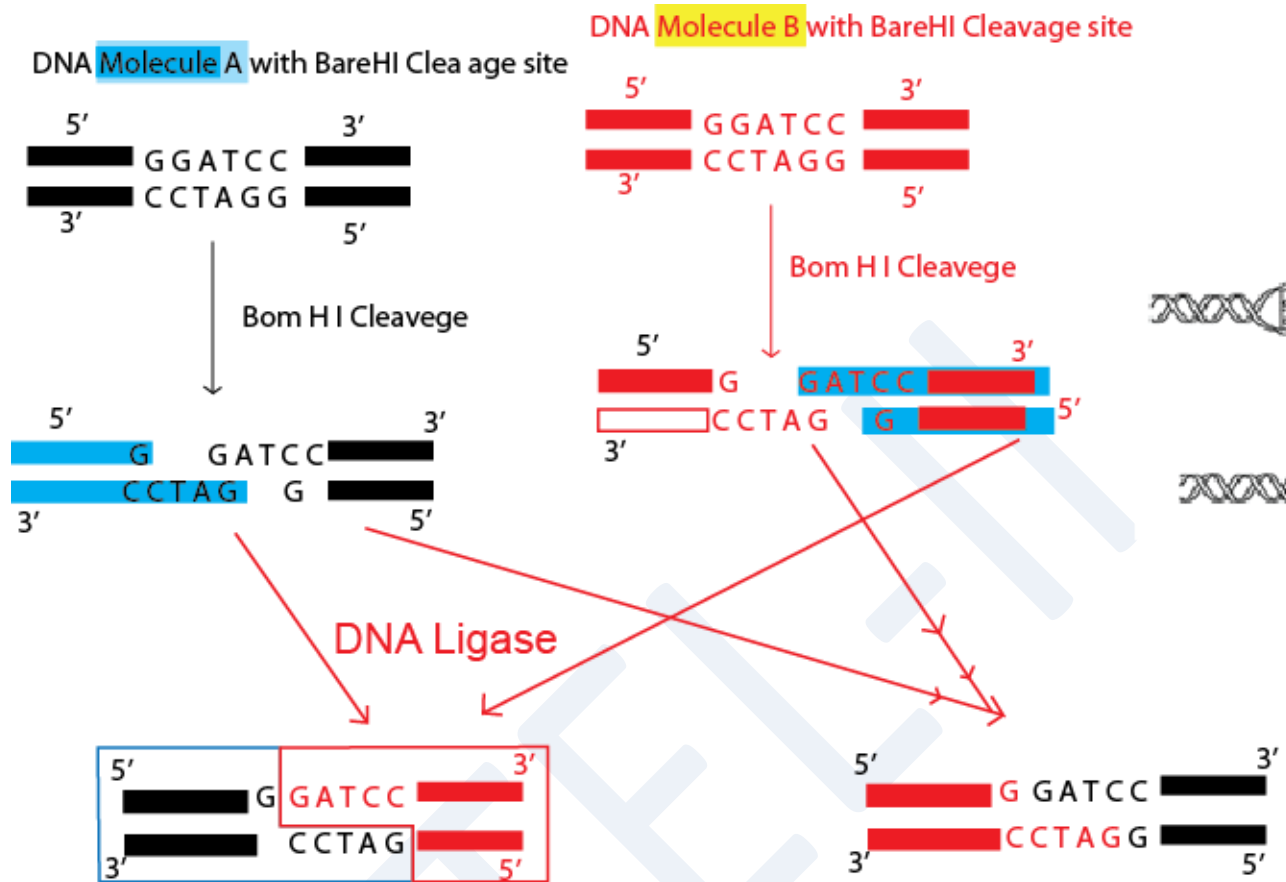**Figure 2:** Cleavage sites of restriction endonucleases

Figure 3: Two DNA molecules with cleavage site for same restriction enzyme may be combined by ligation reaction.

DNA ligase:

DNA ligases are specific type of enzymes that are involved in DNA repair mechanism in the cell. DNA ligases close the nicks in the DNA by forming phosphodiester bond between the 5′-phosphate of a nucleotide on one fragment of DNA and the 3′-hydroxyl of another (Fig. 3). Inside the cell DNA ligases are mainly involved in DNA repair and in the joining of okazaki fragments formed during DNA replication. DNA ligases are mainly classified into two classes. The first class of ligases is found only in bacteria which use NAD^+ as a cofactor. The second class of ligase is mainly found in eukaryotes, viruses and bacteriophages. The second class of ligase use ATP as a cofactor. The most popularly used ligase in molecular biology studies is the smallest known ATP dependent DNA ligase (41 kDa) from bacteriophage T7. DNA ligases from most of the eukaryotes are much larger (>100 kDa) but all DNA ligase are known to share some common sequence and probably structural motif. T4 DNA ligase is known to ligate sticky ends

of DNA more efficiently and but is less efficient in ligating blunt ends. Generally a high concentration of the enzyme is required *in vitro* reactions. To increase the reaction efficiency, the blunt ends are often modified using the enzyme terminal deoxynucleotidyl transferase which adds poly (dA) to DNA fragment from one source and poly (dT) to the DNA from another source. The complementary tails can form hydrogen bond and get ligated to generate a recombinant DNA.