Lecture 10: DNA Libraries (PART-I)

Introduction-Gene sequence are arranged in genome in a randon fashion and selecting or isolating a gene is a big task especially when the genomic sequences are not known. A small portion of genome is transcribed to give mRNA where as a major portion remained untranscribed. Hence, there are two ways to represent a genomic sequence information into the multiple small fragments in the form of a library: (1) Genomic library (2) cDNA library.

Preparation of Genomic Library-A genomic library represents complete genome in multiple clones containing small DNA fragments. Depending upon organism and size of genome, this library is either prepared in a bacterial vector (discussed later in future lectures) or in yeast artificial chromosome (YAC). An outline of the construction of genomic library is given in Figure 10.1. it has following steps:

1. Isolation of genomic DNA

2. Generation of suitable size DNA fragments

3. Cloning in suitable vector system (depending on size)

4. Transformation in suitable host .

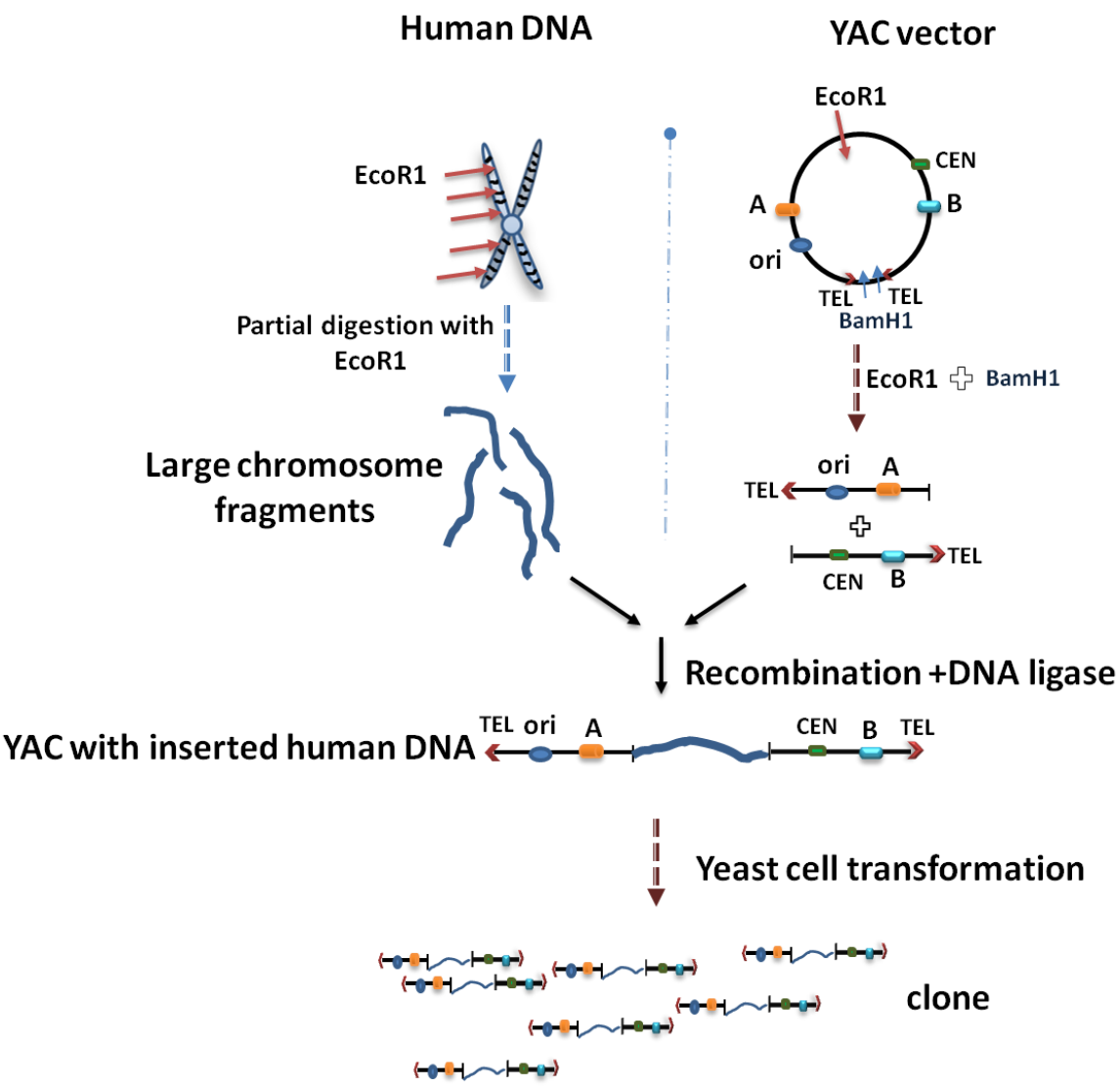

Figure 10.1: Contruction of Genomic library.

1. Isolation of genomic DNA- Isolation of genomic DNA has following steps:

1. Lysis of cells with detergent containing lysis buffer.

2. Incubation of cells with digestion buffer containing protease-K, SDS to release genomic DNA from DNA-protein complex.

3. Isolation of genomic DNA by absolute alchol precipitation.

4. Purification of genomic DNA with phenol:chloroform mixture. Chloroform:phenol mixture has two phases, aquous phase and organic phase. In this step, phenol denatures the remaining proteins and keep the protein in the organic phase.

5. Genomic DNA present in aqueous phase is again precipitated with absolute alchol.

6. Genomic DNA is analyzed on 0.8% agarose gel and a good prepration of genomic DNA give an intact band with no visible smear (Figure 10.2).

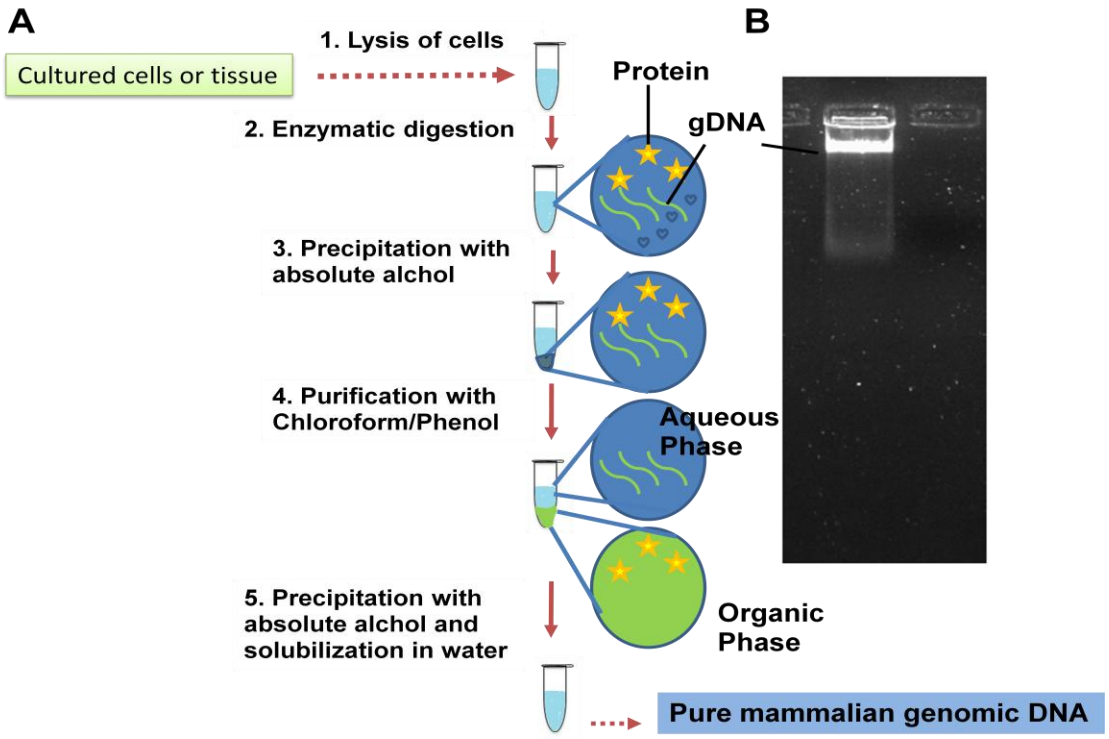

Figure 10.2: Genomic DNA isolation. (A) Different steps in genomic DNA isolation. (B) Agarose gel analysis of isolated genomic DNA.

2. Generation of suitable size fragments- Next step generation of genomic DNA into suitable small size fragments.

Restriction digestion: Genomic DNA can be digested with a frequent DNA cutting enzyme such as EcoR-I, BamH-I or sau3a to generate the random sizes of DNA fragments. The criteria to choose the restriction enzyme or pair of enzymes in such a way so that a reasonable size DNA fragment will be generated. As fragments are randomly generated and are relatively big enough, it is likely that each and every genomic sequence is presented in the pool. As size of the DNA fragment is large, complete genome will be presented in very few number of clones. In addition, genomic DNA can be fragmented using a mechanical shearing.

If a organism has a genome size of $2x10^7$ kb and an average size of the fragment is 20kb, then no. of fragment, n= 10^6. In reality, this is the minimum number to represent a given fragment in the library where as the actual number is much larger.

The probability (P) of finding a particular genomic sequence in a random library of N independent clone is as follows:

$N=\ln (1-P)/\ln (1-1/n)$……..**(Eq 10.1)**

Where, N=number of clones, P=probability, n= size of average fragment size

3. Cloning into the suitable vector-The suitable vector to prepare the genomic library can be selected based on size of the fragment of genomic DNA and carrying capacity of the vector (Table 10.1). Size of average fragment can be calculated from the Eq 10.1 and accordingly a suitable vector can be choosen. In the case of fragment generated by restriction enzyme, vector can be digested with the same enzyme and put for ligation to get clone. In the case of mechanical shearing mediated fragment generation, putting these fragment needs additional effort. In one of the approachs, a adopter molecule can be used to generate sticky ends, alternatively a endonuclease can be used to generate sticky ends.

4. Transformation to get colonies- Post ligation, clones are transformed in a suitable host to get colonies. A suitable host can be a bacterial strain or yeast. Different methods of delivering clone into the host cell is discussed in future lectures.

Table 10.1: Carrying capacity of different vectors

S.NO	Vector	Insert Size (MB)
1	Plasmids	15
2	Phage lambda	25
3	Cosmids	45
4	Bacteriophage	70-100
5	Bacterial artificial chromosome (BAC)	120-300
6	Yeast artificial chromosome (YAC)	250-2000

Quiz

Q1: The size of mouse genome is $5.6x10^6$ Kb and average cloned fragment size is 40kb. How many minimum number of clones are required to represent a particular sequence ?

Ans: $1.4x10^5$

Q2: In Question no. 1, How many clones are needed to represent a sequence with a probability of 95% ?

Ans: $4.2x10^5$

Q3: What are the different methods of generating random genomic fragments ?

Ans: Partial Restriction digestion and mechanical shearing.

Q4: Which vector is used to create genomic library for human genome in genome sequencing project?

Ans: Bacterial artificial chromosome and Yeast artificial chromosome.

Q5: What is the limitiation of genomic DNA library?

Ans: it contains sequence with no information of gene expression status.

Lecture 11: DNA Libraries (PART-II)

Introduction- In the previous lecture we discussed construction of genomic library and now we will discuss construction of cDNA library.

Contruction of cDNA library-A cDNA library represents mRNA population present at a particular stage in a organism into multiple clones containing small DNA fragments. An outline of the construction of cDNA library is given in Figure 11.1. it has following steps:

1. Isolation of mRNA

2. Preparation of complementary DNA fragments-

3. cloning in suitable vector system

4. Transformation in suitable host .

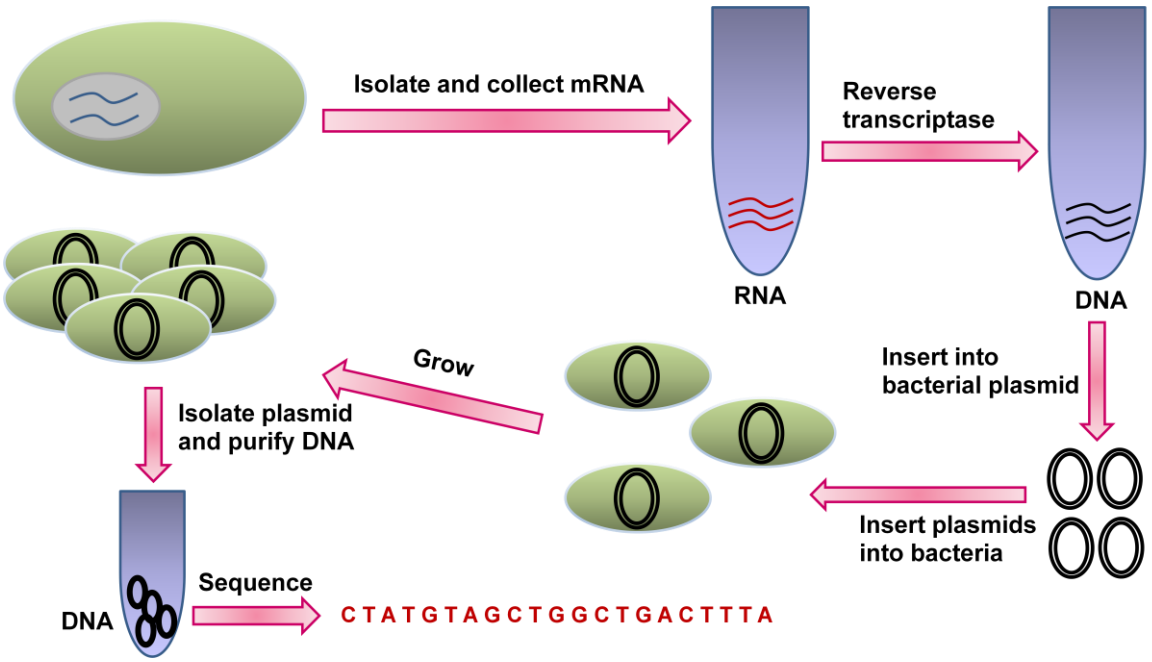

Figure 11.1: Steps in construction of cDNA library.

1. Isolation of mRNA- The structure of a typical mRNA is given in Figure 11.2. it has a CAP structure at 5', coding sequence and a poly A tail at its 3' region. The Nucleotide A forms 2 hydrogen bonding with nucleotide T and this pairing is very specific. Exploting this feature, m-RNA population can be isolated from RNA pool using a poly-T affinity column. The Steps in m-RNA isolation from cell is given in Figure 11.3. it has following steps:

1. Release of total RNA either by a lysis buffer containing detergent or by homogenization in the case of hard tissue.

Figure 11.2: Structure of a typical mRNA.

2. Mixing of poly-T containing beads with the total RNA prepration. Due to mutual exclusive affinity, mRNA binds to the poly-T beads.

3. Wash the beads with washing buffer to remove non-specific cross contaminating species.

4. Elute the mRNA from beads; its purity can be checked on polyacryalamide gel (discussed later in future lectures).

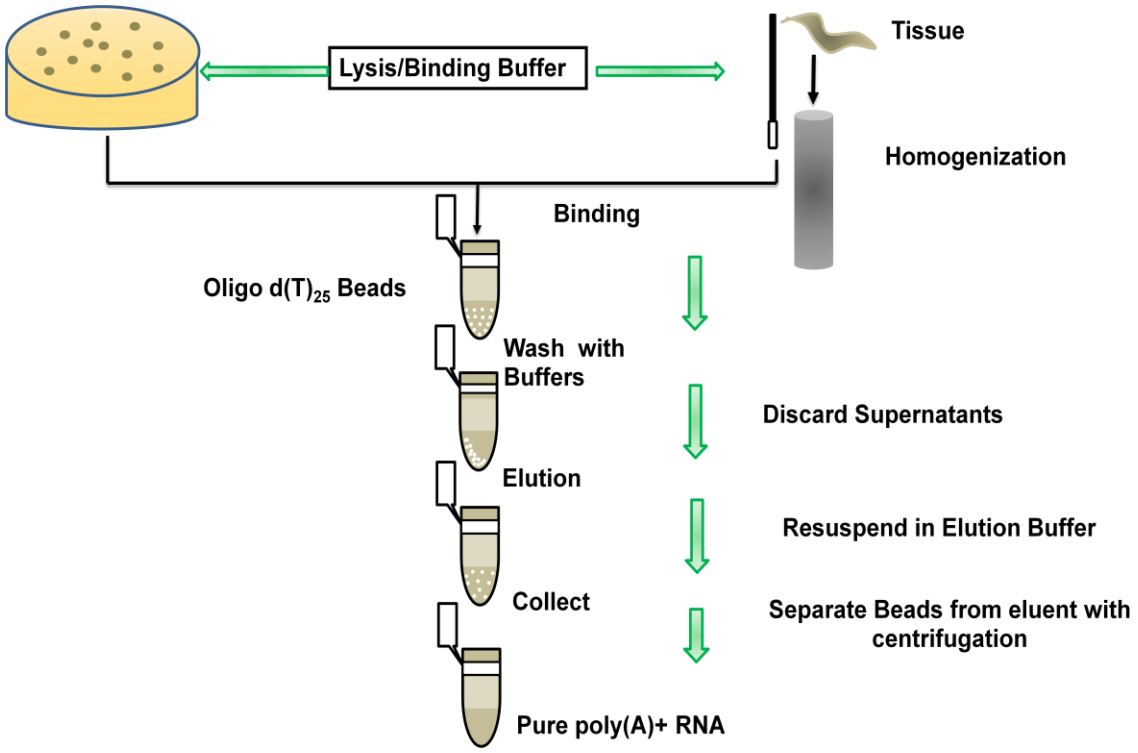

Figure 11.3: Isolation of mRNA from total cell lysate.

Preparation of complementary DNA (cDNA)- Multiple approaches have been developed to prepare complementary DNA (cDNA) from isolated mRNA. In all approaches the 3 steps are performed.

1. First strand synthesis with a reverse transcriptase.

2. Removal of RNA template

3. Second strand synthesis

Homopolymer tailing- This method exploits the presence of poly A tail present on m-RNA to synthesize first DNA strand followed by degradation of RNA template and synthesis of second strand. A schematic is given in Figure 11.4 to show representative developed method. It has following steps-

1. An oligo dT primer is used with mRNA as template to prepare first strand of DNA with the help of reverse transcriptase and dNTPs.

2. After the synthesis of first strand, terminal transferase is used to add C nucleotides on 3'of both mRNA and newly synthesized firsr strand of DNA.

3. DNA: RNA hybrid is loaded on a alkaline sucrose gradient. This step will hydrolyze RNA and allow the full recovery of cDNA.

4. Next, an oligo dG primer is used with cDNA as template to prepare second strand of DNA with the help of reverse transcriptase and dNTPs.

Gubber-Hoffman method-The method of Gubber-Hoffman is shown in Figure 11.5. In this approach, after first strand synthesis using oligo dT primer in the presence of reverse transcriptase and dNTPs. DNA:RNA hydrid is treated with RNase H to produce nicks at multiple sites. Then DNA polymerease is used to perform DNA synthesis using multiple fragment of RNA as primer to synthesize new DNA strand. This method produces blunt end duplex DNA product.

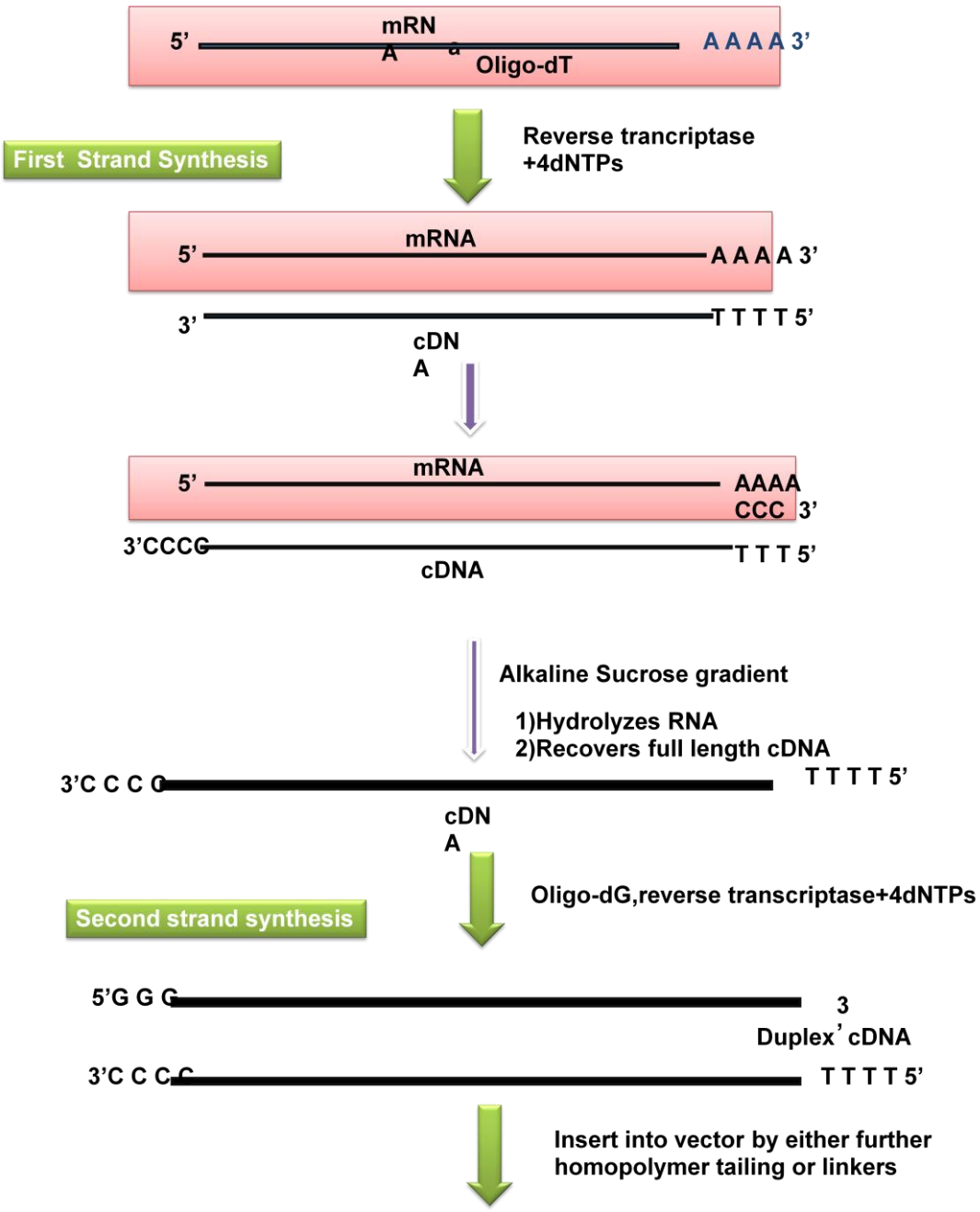

Figure 11.4: Generation of complementary DNA from m-RNA by homopolymer tailing method.

3. Cloning of cDNA into the vector-The cDNA is ligated into the suitable vector to generate clone.

4. Transformation to get colonies- Post ligation, clones are transformed in a suitable host to get colonies. A suitable host can be bacterial strain or yeast. Different methods of delivering clone into the host cell is discussed in future lectures.

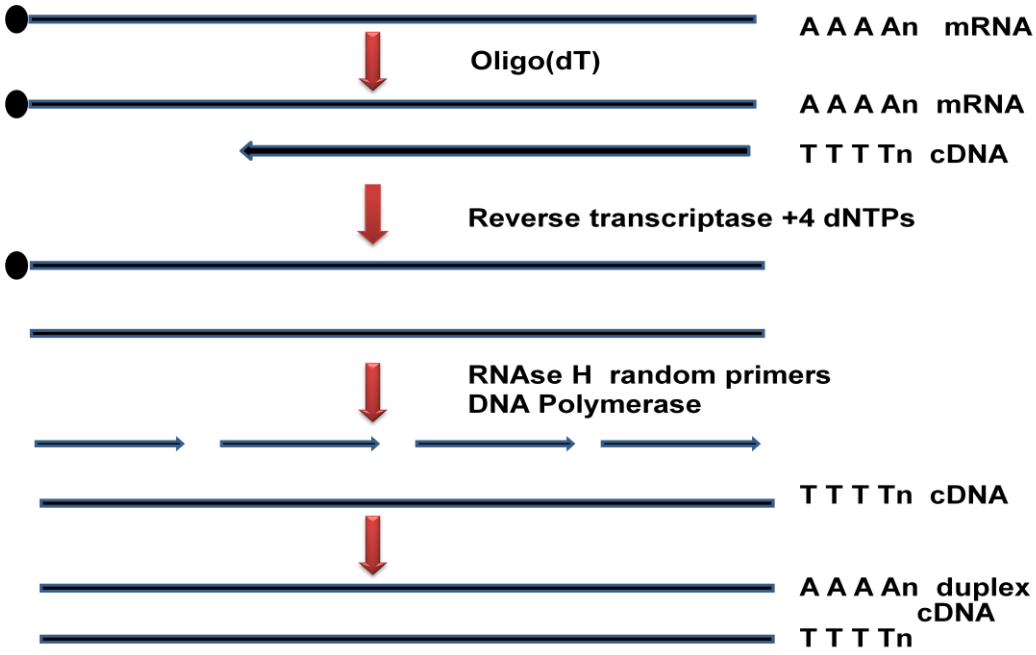

Figure 11.5: Generation of complementary DNA from m-RNA by Gubber-Hoffman method.

Quiz

Q1:cDNA library represents the part of genome which gives……

Q2: Oligo dT primer is used for synthesis of first DNA strand because…….

Q3: Discuss the different steps of construction of cDNA library ?

Hint: please follow the lecture and try to answer this question.

Q4: What is the advantage of gubbler-hoffman method over other method of cDNA synthesis?

Hint: please follow the lecture and try to answer this question.

Q5: What are the limitiations of cDNA library?

Lecture 12: Identification and Isolation of a gene

Introduction-Identification and isolation of a particular gene is essential for development of the biotechnology applications. With the availability of genomic sequences the task is getting easier day-by-day. The approaches we would like to discuss today is to identify and isolate a gene fragment from an organism. In previous two lecture we have discussed presenting genomic DNA either not associated with the production of protein (**Genomic Library**) or responsible for production of protein (**cDNA library**). Now in the present lecture we will discuss screening and isolation of gene with known structural (DNA sequence) or functional attributes (enzyme activity or particular antigenic epitope).

There are 3 different searchable criteria to identify a particular gene from an organism:

1. DNA sequence-This properties can be use to search both genomic library and cDNA library to identify the gene.

2. Expression of a particular protein with immunogenic epitope-This property can be partially useful to screen genomic library due to truncation of a full gene or no expression of a gene fragment. But this approach suits well to screen cDNA clones.

3. Enzymatic activity- This property exploits the ability of a protein fragment to exhibits enzymatic activity. It is useful for the screening of cDNA library but not much for genomic library.

SCREENING BY DNA HYBRIDIZATION

DNA sequence information can be exploited with a general rule that nucleotide present in a DNA sequence provides a specificity due to unique base pairing preference of nucleotides. “A” is always making base pairing with “T” and “G” is making base pairing with “C”.

As a result a particular DNA sequence can be identified by a complementary single stranded DNA sequence. The DNA sequence used for this purpose is called as “Probe”. After-wards the position of probe can be identified by a suitable detection system. The position of probe is the actual site of desirable clone of containing specific sequence. This complete procedure of colony hybridization is given in Figure 12.1 and it has following steps:

1. Preparation of suitable radioactive probe.

2. Preparation of replica plate

3. Transfer of colonies on nitrocellulose membrane.

4. Hybridization with a specific probe.

5. Washing and development of membrane by autoradiography.

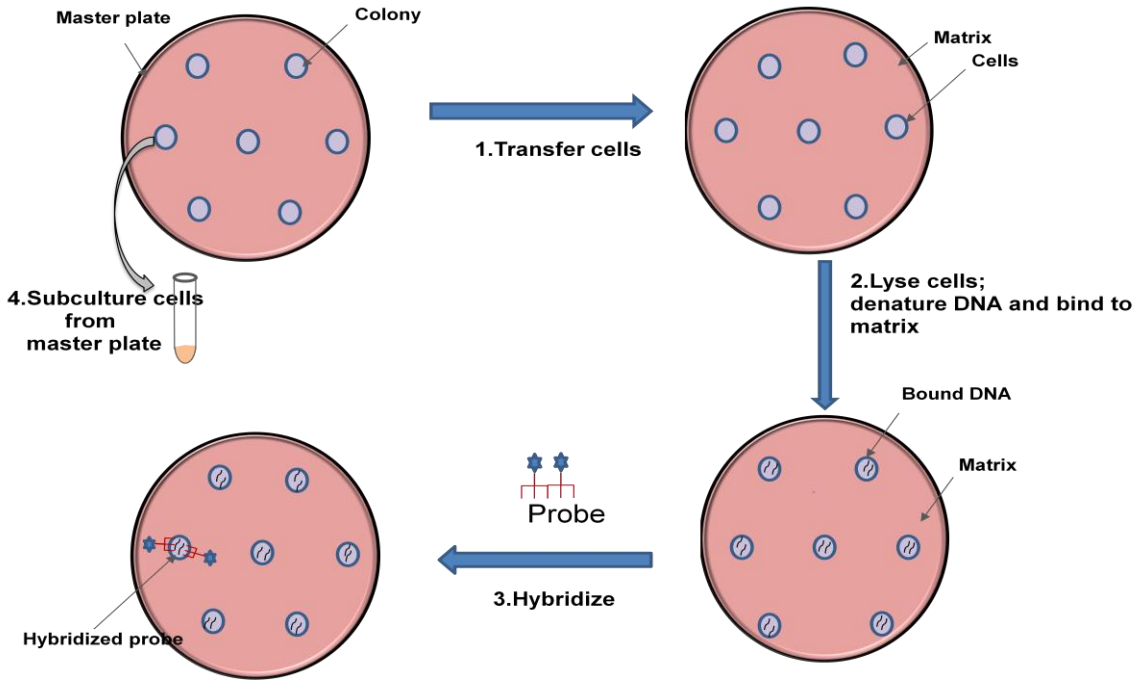

Figure 12.1: Screening a library with a radioactive probe by colony hybridization.

1. Preparation of radioactive probe. There are two different method used to label a single stranded DNA probe either at terminal or throught the sequence.

A. Random primer method- In this method, a random primer is used to anneal to the template and then a PCR reaction is performed in the presence of radiolabeled nucleotide. After PCR, newly synthesized DNA strand is labeled with radioactive nucleotide. The whole process is given in Figure 12.2, and it has following steps-

1. The source double stranded DNA is denatured to generate the single stranded DNA template.

2. A random primer is added and allowed it to anneal to the template strand. It will anneal to the random position through out the sequence at multiple places.

3. Primer will anneal to the template strand and now klenow will start the synthesis of new DNA strand.

4. Newly synthesized DNA will give short stretches of multiple labeled DNA probes.

B. Terminal transferease- In this method, a terminal transferase enzyme will label the probe at the ends to the last nucleotide of the probe (Figure 12.3). Probe is incubated with the labeled nucleotide and terminal transferase enzyme will add the labeled nucleotide at the end. A partial purification with gel filtration column will give labeled primer.

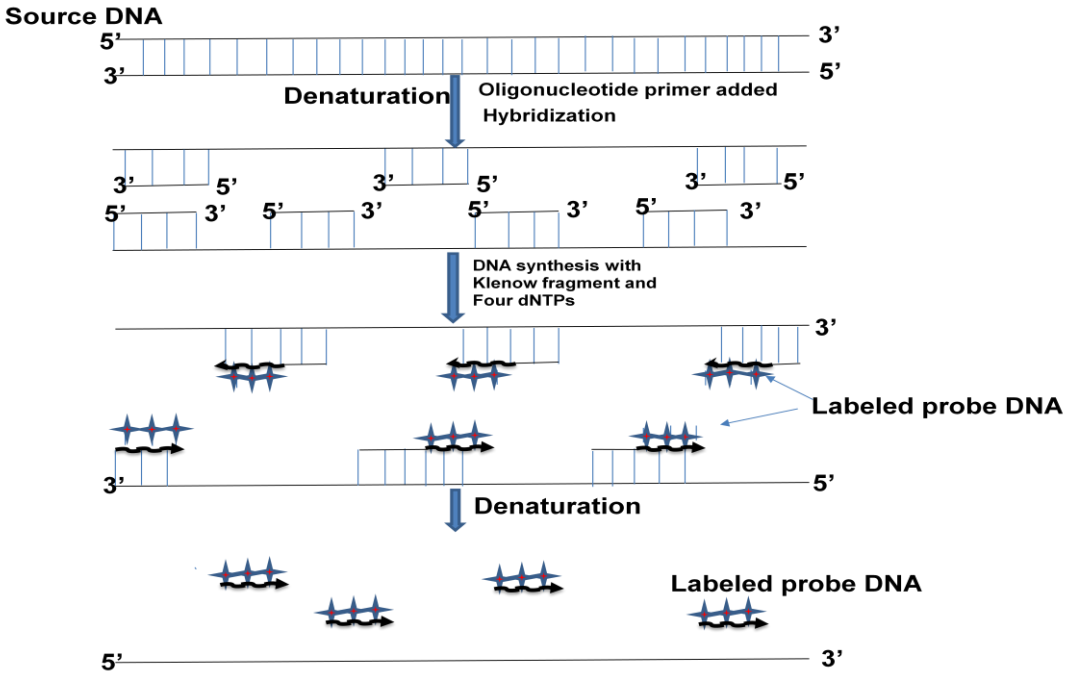

Figure 12.2: Preparation of radioactive probe by random primer method.

2. Preparation of Replica plate- As original genomic or cDNA library is precious and will be consumed in later stage, all procedure is performed with the replica plate containing clones in an identical manner.

3. Blotting-The clone is transferred on a nitrocellulose membrane with retaining identical pattern of colonies on master plate. The cells on the membrane are lysed and released DNA is denatured, deproteinated and allowed to bind the membrane.

3. Hydridization-A labeled probe prepared in step 1 will be added. Probe will binds to the target DNA due to base pairing (Figure 12.1). The membrane is washed to remove unbound probe.

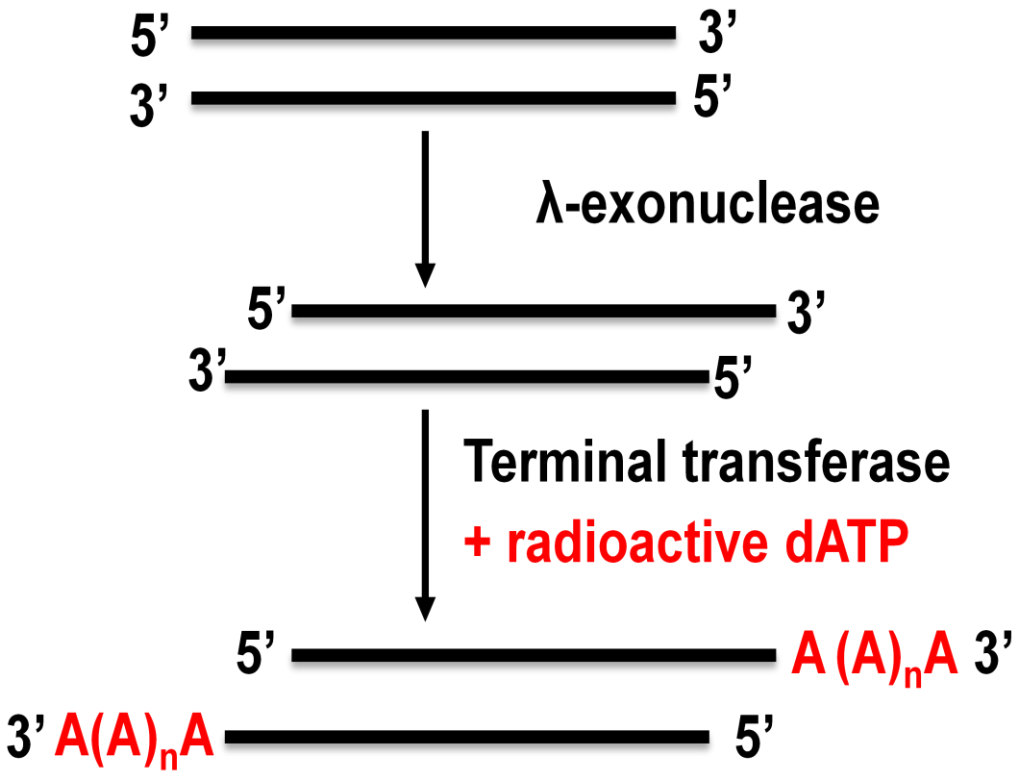

Figure 12.3: Preparation of radioactive probe by terminal transferase method.

4. Development of blot (Autoradiography)-The position of labeled probe is detected by autoradiogram.The position of signal on membrane can be matched with the master plate to get location of corresponding colony.

SCREENING BY IMMUNOLOGICAL METHODS

This method is based on the specificity of antibody towards its antigenic epitope present on the protein expressed in a particular clone (Figure 12.4). A number of disease associated gene have been identified by this method. Due to increased expression or unique expression of a particular protein in a disease condition, patient body develops antibody against it. The developed antibody is available to use to identify the protein expressing clone. This method has following steps:

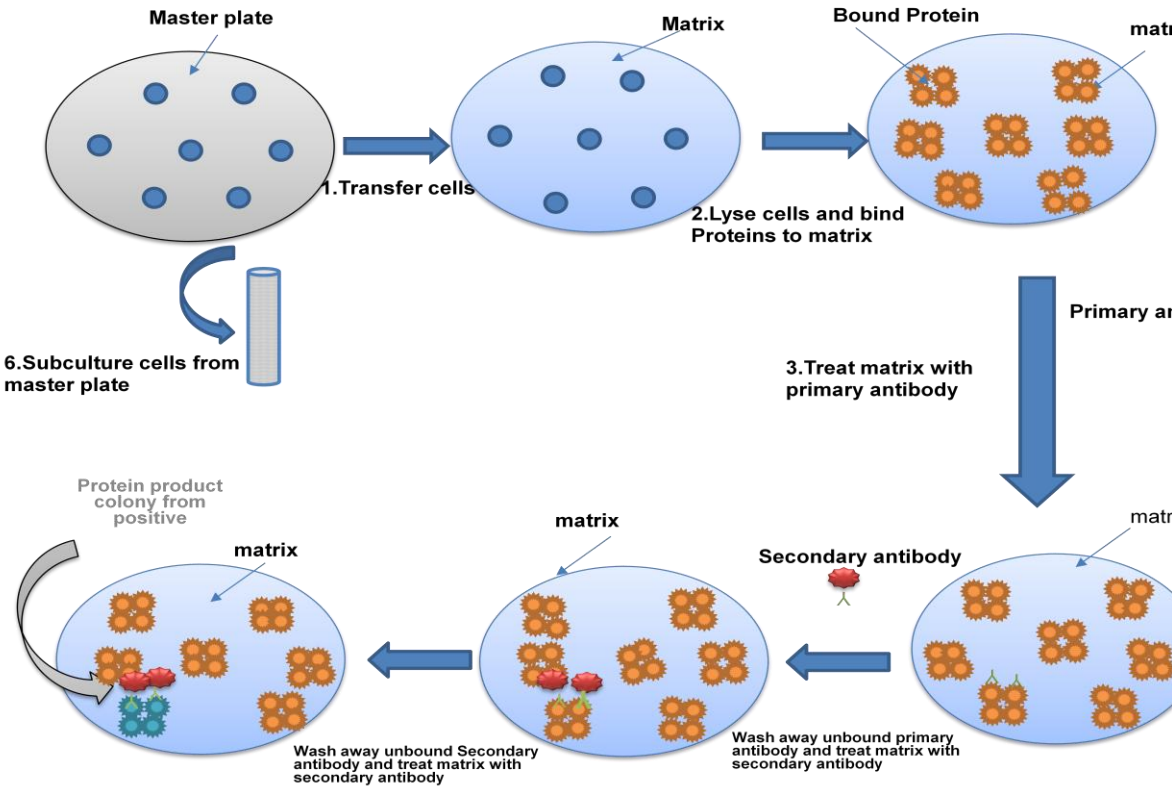

Figure 12.4: Sceening by immunological methods

1. Preparation of Replica plate- As original genomic or cDNA library is precious and will be consumed in later stage, all procedure is performed with the replica plate containing clones in a indentical manner.

2. Blotting-The clone is transferred on a nitrocellulose membrane to get similar pattern of colonies on master plate. The cells on the membrane are lysed and released protein is denatured, and allowed to bind the membrane.

3. Treatment with primary antibody-The membrane is incubated with antibody having immunoreactivity towards a particular protein. The primary antibody will binds to the target protein due to exclusive specificity towards antigen (Figure 12.4). The membrane is washed to remove unbound primary antibody.

4. Treatment with secondary antibody-The membrane is incubated with secondary antibody recognizing primary antibody. Secondary antibody is labeled with an enzyme (Horse raddish peroxidase or alkaline phosphatase) to use to give readable signal. The secondary antibody will binds to the primary antibody and will allow to detect the location of primary antibody. The membrane is washed to remove unbound secondary antibody.

THINK TANK??

Why enzyme labeled secondary antibody is used instead of labeled primary antibody ?

Development of blot-The position of secondary antibody is detected by performing enzymatic activity.The position of signal on membrane can be matched with the master plate to get location of corresponding colony.

SCREENING BY ENZYMATIC ACTIVITY

This method is based on the ability of protein to exhibit an enzymatic activity. This method is not very specific but allow us to identify a class of protein with known enzymatic activity.

Isolation of gene-Once the position of a clone is known, it is extracted from the master plate and plasmid is isolated. In few cases, clone is further diluted to check the homogeneity of clone. The purity of the clone and presence of clone is further tested with a PCR using sequence specific primers.

HOME ASSIGNMENT

The gene fragment with the nucleotide seqeunce as given below needs to be identify from the human genome. Give the complete strategy to answer the following questions (draw images as required) :

1. Identify the human gene and comparative study of expression of this gene in liver and muscles ?

2. Presence or absence of intronic regions ?

3. Position of this gene on human chromosomes ?

Nucleotide Sequence:

GCATGTGAATCCCAGTAGTTAAACATGACTGGTTGACCCAGGCCTTGCAGGAGCCCAGAGGCAAGGAGGG
AATGTGGGTGAGCCGCAGGAGCTTAAGCTAGTCAAAGTCAAGTTCAGAGGCCAGTGGACCTGCACAAACT
TGAAAAAGGCTCAACACCTGTTGTAAATATTTGCTTGCTACTGAGTCCTACTCAAGCCATCTCCATATCA
CCATGAATGAAGTCAGACTTTTAGGGAAAAGTGGTTAACAATCTGGATTTCTGCGGTAGCAGGGTGGATC
ATATAATGCTTACTCCTGACACCACGCCTGACACCTATTCAGTTTCTCAGCTTTACTCAAAGCCTCCATG
ACTTAGCAGGTTCAGGTTAGGCGACCTTCTTAGGTAATCCTAATAACAGCTGGCCTTCCCTTCTCTCAGC
ACTTTTTCCCCGTGTTATTGTCTGATTATTTGTCTATGTCCACGTCCTTTGCTGTTGTACTCCCAGAGCC
CAATAGGCAGGGGGCCCAGGACACAGCAGACATCCAGTATATTGGATGACTCCCATGGGTTAGTGGTTAT
GGCTGAACATAACATGAGAGGCCTTCTCTCTCAGAAGAATGTGATGAAATATGCAATTGATGAACATCCA
GAATCCAGATATTCACCTGTCTTGAAAATTGCACTTCCAATCCACACTCTCAGTGCCTGTGACAGTATAA
ACTAGCTCGAGCATTGCAGAAAACAATCAACAACATGTCACAAGAACCTTAAAAAGTTCAAATGACTTGA
ACCAAAAATGACCTCTGAAAACAGGTCCTAAGGAAATAAATACAAATGTAGAGGAAGGATGTCACTGTAC

Question : What are the methods can be use to identify the gene from human genome ?

Lecture 13: Basics of Cloning (PART-I)

Over-view of cloning: Cloning refers to the process of producing genetically identical organism. In general, for biotechnology related applications, cloning is used to produce DNA, either as a part of a functional gene or part of regulatory region such as promoter. An outline of basic steps involved in cloning is given in Figure 13.1. It has multiple steps to achieve cloned gene in a vector for amplification.

1. Isolation or amplification of gene fragment from genome of the organism.

2. Restriction digestion of gene fragment and vector

3. Ligation of cut gene product and vector

4. Insertion of ligated DNA or recombinant DNA into the host.

5. Screening and selection of cells containing recombinant DNA.

A complete procedure of cloning involves usage of multiple enzyme or molecular tools to perform these steps. Before discussing minor details of these procedure, we will discuss the enzymes involved in cloning procedure and their enzymatic mechanism. Besides this we will also discuss about molecular accessories to facilitate cloning in special conditions.

Molecular Accessories Developed for cloning

1. Linker Molecule-An amplified foreign DNA may have restriction enzymes at their terminus to give cohesive end to facilitate ligation into the vector. But in cases when foreign DNA is a genome product and there is least possibility to keep restriction enzymes at the ends. Cloning of these fragments is facilitated by the help of a linker molecules. Linker molecules are the short double stranded DNA (8-10 nucleotide long) and has restriction sites at their ends. Forex, a typical linker molecule with EcoRI site is shown in Figure 13.2. Linker molecule is incubated with foreign DNA and ligated by the action of T4 DNA ligase to generate chimeric DNA. The chimeric DNA is digested with EcoRI to generate cohesive ends. It is now incubated with EcoRI digested vector in the presence of DNA ligase to get circular clone (Figure 13.2).

2. Adaptor Molecule- During a restriction digestion of chimeric DNA, in few cases the restriction enzyme cuts the linker molecule or foreign DNA. This can be avoided by choosing different restriction enzymes but in some cases

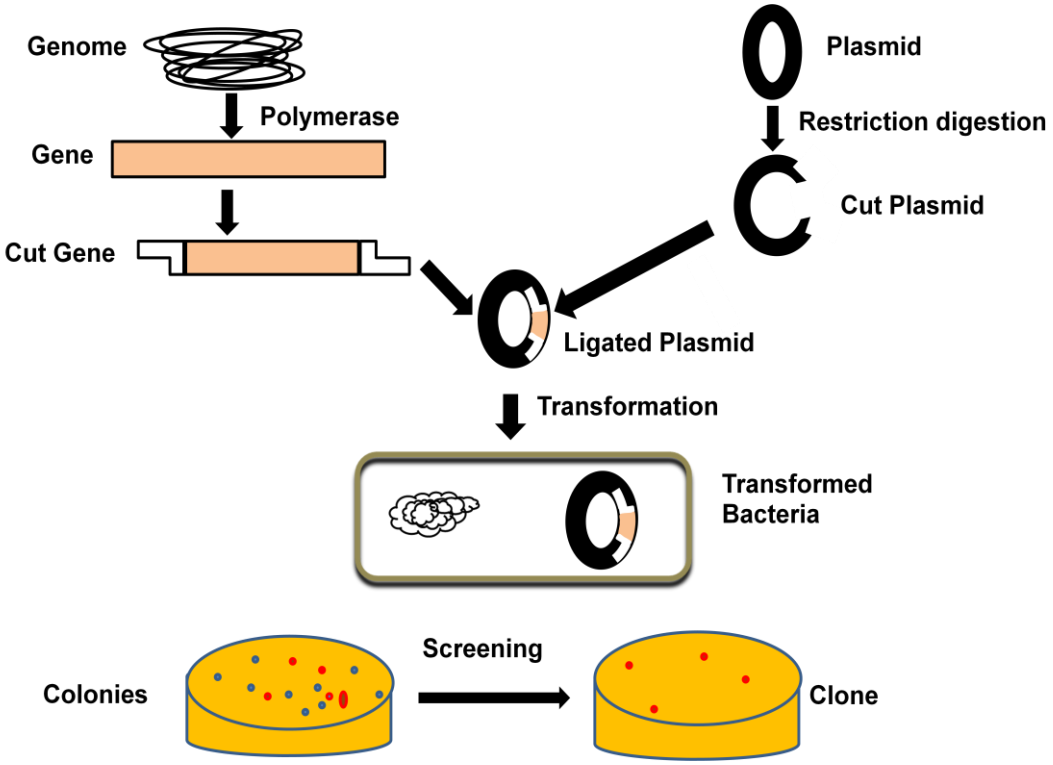

Figure 13.1: An Over-view of different steps involved in cloning.

it is not possible. In such conditions, an adopter molecule can be added. Adaptor molecules are 8-10 nucleotide long double stranded molecule with flanking DNA sequence to provide cohesive ends (Figure 13.3). These cohesive ends have free hydroxyl group to facilite efficient ligation to the vector. Chimeric DNA containing adaptor molecule is incubated with BamH-I digested vector in the presence of DNA ligase to get circular clone (Figure 13.3).

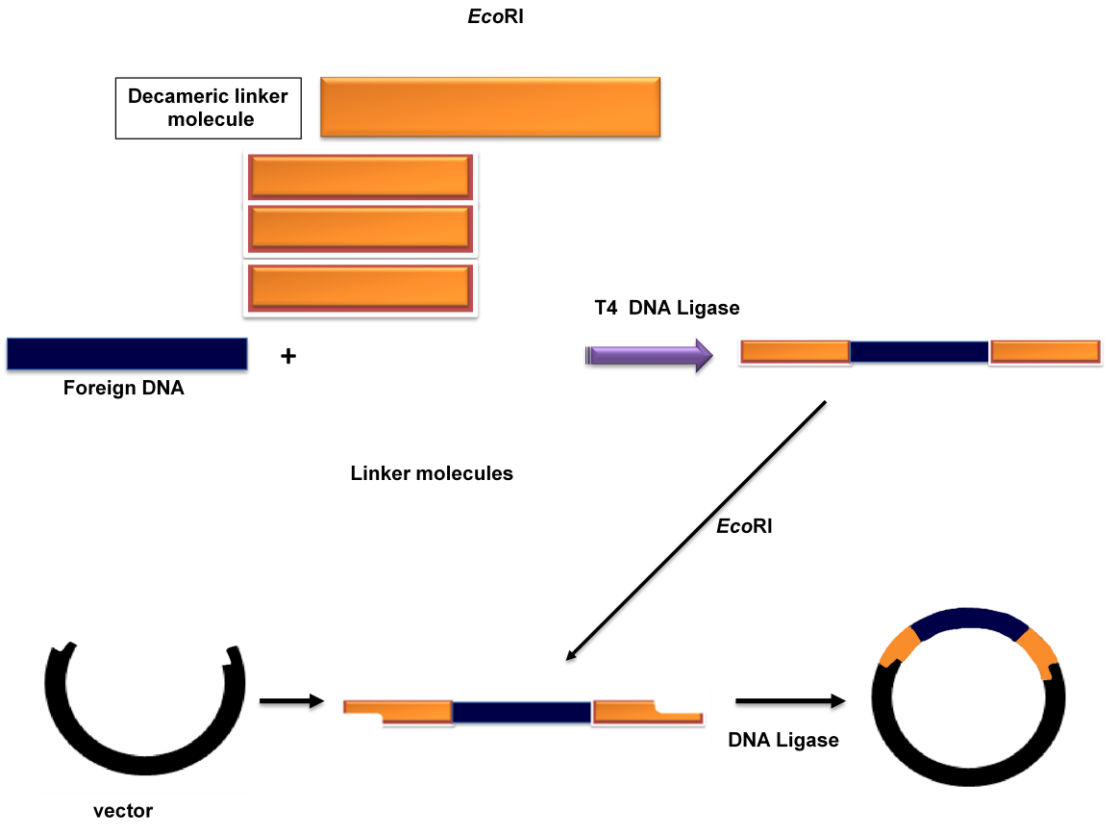

Figure 13.2 : Linker molecule and its usage in cloning of a blunt DNA.

Enzymes used in molecular cloning

1. Polymerase- The first requirement of a molecular cloning to produce or isolate DNA of interest from the genomic sequence. One of the easiest process to achieve it is by amplifying the foreign DNA with site specific primers. A details of these steps will be discussed in future lecture.

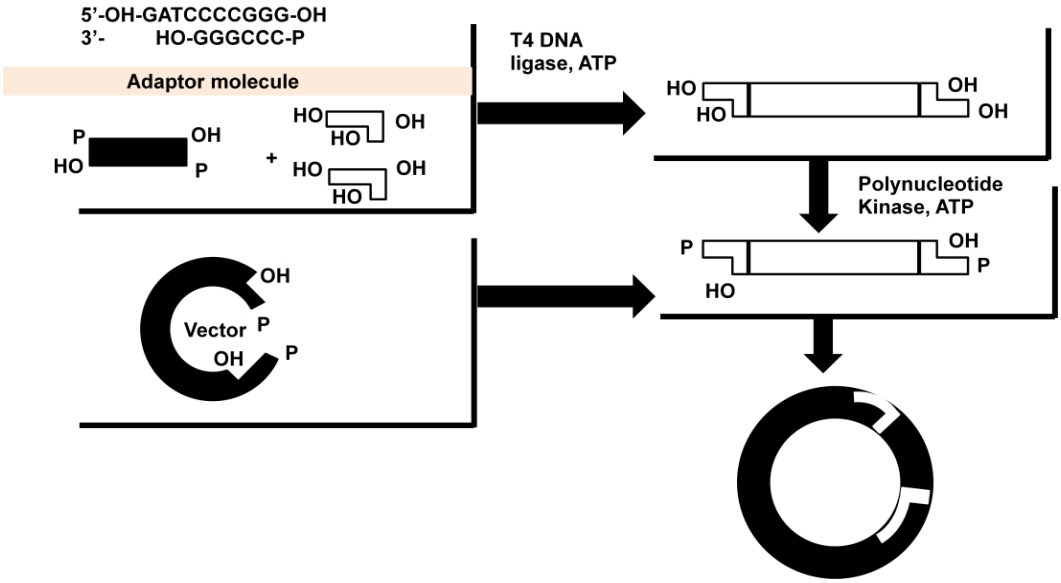

Figure 13.3 : Adaptor molecule and its usage in cloning.

2. Ligase-Joining two DNA to generate the chimeric DNA is the basis of cloning. As discussed in Figure 13.1, it is an essential step to generate clone containing foreign DNA in a vector. When cohesive ends are generated by the action of restriction endonuclease on DNA associated with each other, a nick remains to be sealed and give complete circular DNA. **What DNA ligase is doing?** It is an enzyme requires ATP or NAD^+ as a cofactor to catalyze ligation reaction. Ligase is processing ATP to generate AMP, and then AMP is making adduct with enzyme to form ligase-AMP complex. This complex is binding to the 3' and 5' of DNA bearing nick and bringing them together. AMP is released and phosphodieaster linkage is formed between 3' and 5' end to seal the nick (Figure 13.4).

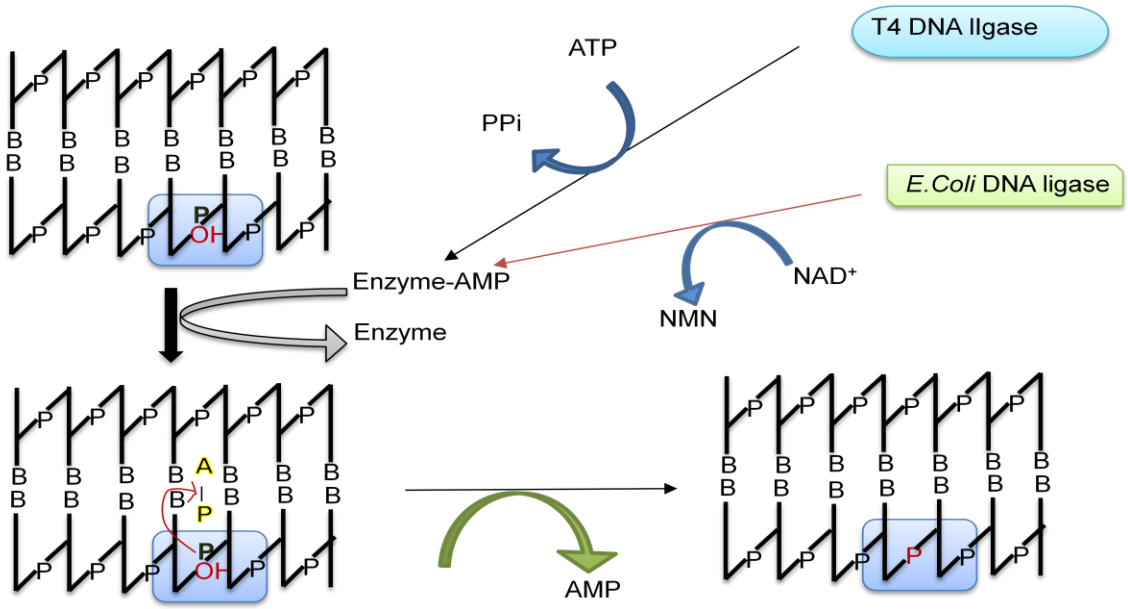

Figure 13.4 : Mechanism of DNA ligase.

3. Alkalike phosphatase- Digested linear plasmid containing cohesive ends on both sites with phosphate has a tendency to recircularize (Figure 13.5). Removing terminal phosphate group prevents this possibility and for this purpose, alkaline phosphatase is used. Alkalike phosphatase removes 5'-terminal phosphate groups and in this condition, only in the presence of insert DNA as it will supply phosphate group at both ends to facilitate the ligation reaction (Figure 13.5).

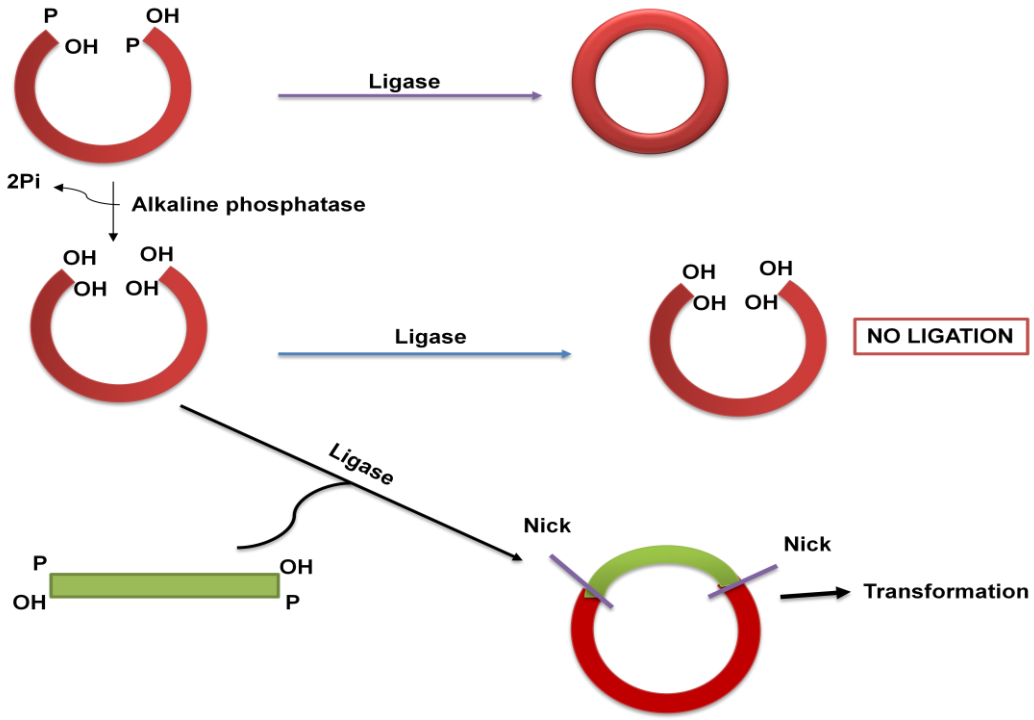

Figure 13.5 : Usage of Alkaline phosphatase in cloning.

Lecture 13: Basics of Cloning (PART-I)

Over-view of cloning: Cloning refers to the process of producing genetically identical organism. In general, for biotechnology related applications, cloning is used to produce DNA, either as a part of a functional gene or part of regulatory region such as promoter. An outline of basic steps involved in cloning is given in Figure 13.1. It has multiple steps to achieve cloned gene in a vector for amplification.

1. Isolation or amplification of gene fragment from genome of the organism.

2. Restriction digestion of gene fragment and vector

3. Ligation of cut gene product and vector

4. Insertion of ligated DNA or recombinant DNA into the host.

5. Screening and selection of cells containing recombinant DNA.

A complete procedure of cloning involves usage of multiple enzyme or molecular tools to perform these steps. Before discussing minor details of these procedure, we will discuss the enzymes involved in cloning procedure and their enzymatic mechanism. Besides this we will also discuss about molecular accessories to facilitate cloning in special conditions.

Molecular Accessories Developed for cloning

1. Linker Molecule-An amplified foreign DNA may have restriction enzymes at their terminus to give cohesive end to facilitate ligation into the vector. But in cases when foreign DNA is a genome product and there is least possibility to keep restriction enzymes at the ends. Cloning of these fragments is facilitated by the help of a linker molecules. Linker molecules are the short double stranded DNA (8-10 nucleotide long) and has restriction sites at their ends. Forex, a typical linker molecule with EcoRI site is shown in Figure 13.2. Linker molecule is incubated with foreign DNA and ligated by the action of T4 DNA ligase to generate chimeric DNA. The chimeric DNA is digested with EcoRI to generate cohesive ends. It is now incubated with EcoRI digested vector in the presence of DNA ligase to get circular clone (Figure 13.2).

2. Adaptor Molecule- During a restriction digestion of chimeric DNA, in few cases the restriction enzyme cuts the linker molecule or foreign DNA. This can be avoided by choosing different restriction enzymes but in some cases

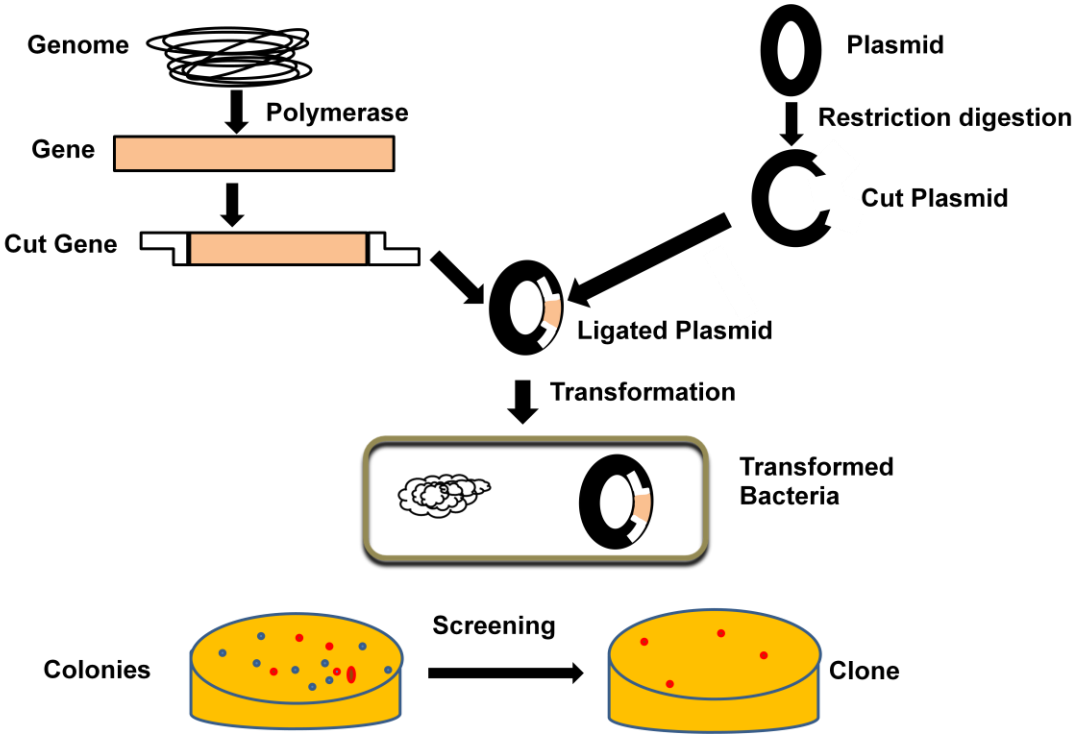

Figure 13.1: An Over-view of different steps involved in cloning.

it is not possible. In such conditions, an adopter molecule can be added. Adaptor molecules are 8-10 nucleotide long double stranded molecule with flanking DNA sequence to provide cohesive ends (Figure 13.3). These cohesive ends have free hydroxyl group to facilite efficient ligation to the vector. Chimeric DNA containing adaptor molecule is incubated with BamH-I digested vector in the presence of DNA ligase to get circular clone (Figure 13.3).

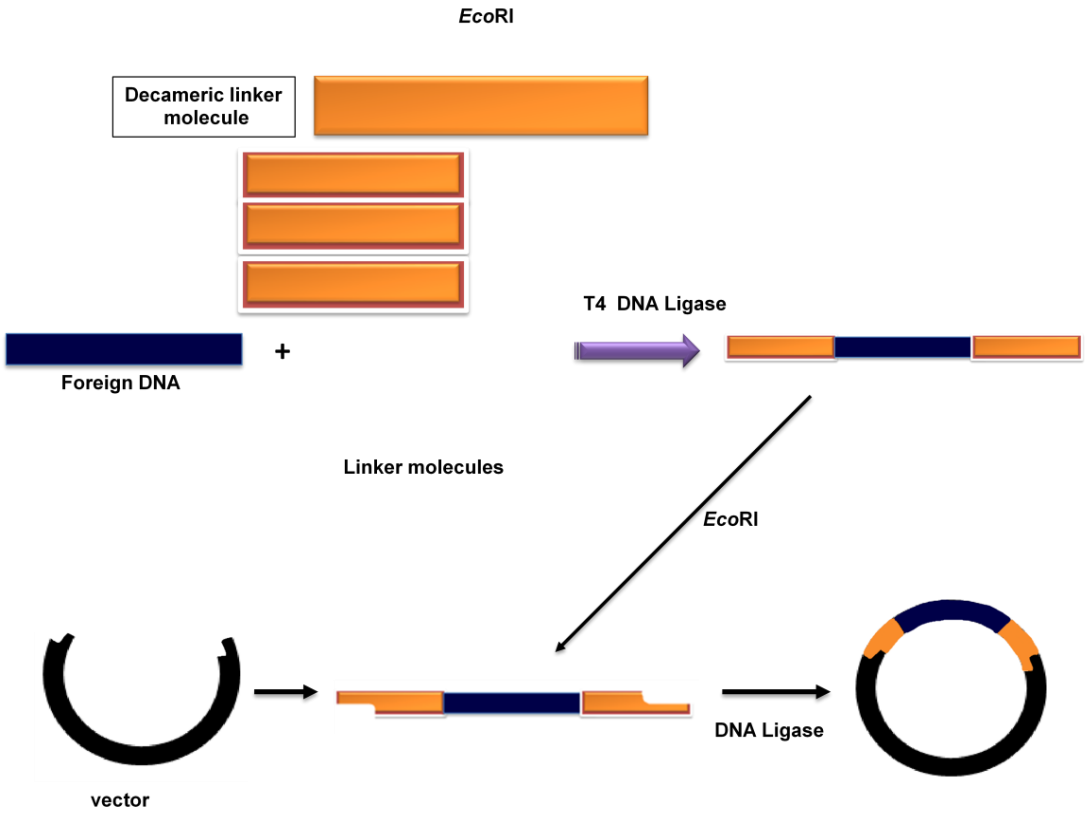

Figure 13.2 : Linker molecule and its usage in cloning of a blunt DNA.

Enzymes used in molecular cloning

1. Polymerase- The first requirement of a molecular cloning to produce or isolate DNA of interest from the genomic sequence. One of the easiest process to achieve it is by amplifying the foreign DNA with site specific primers. A details of these steps will be discussed in future lecture.

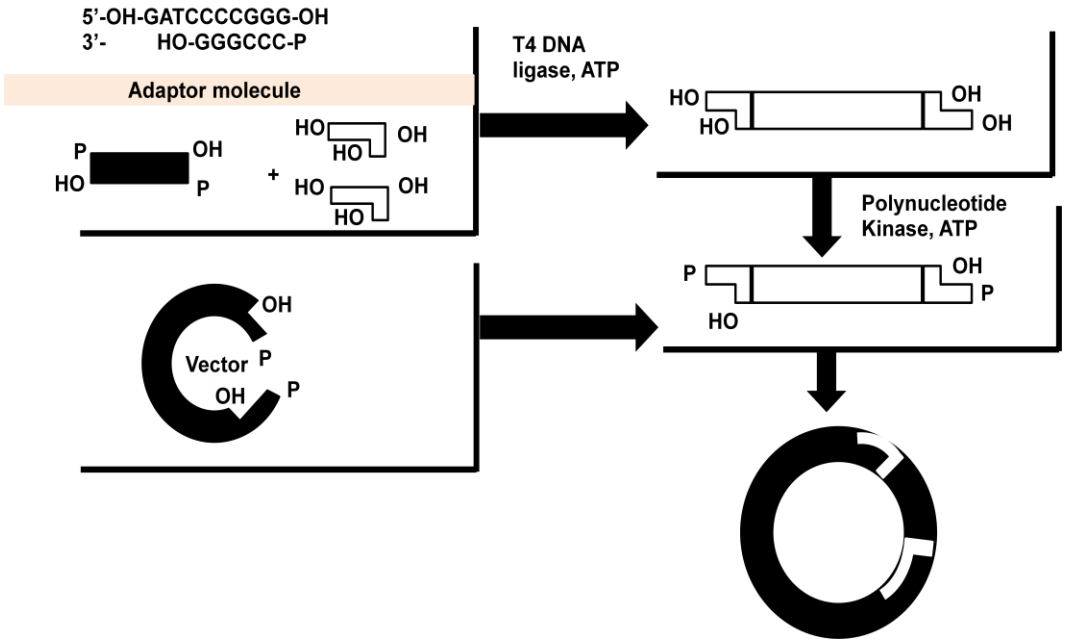

Figure 13.3 : Adaptor molecule and its usage in cloning.

2. Ligase-Joining two DNA to generate the chimeric DNA is the basis of cloning. As discussed in Figure 13.1, it is an essential step to generate clone containing foreign DNA in a vector. When cohesive ends are generated by the action of restriction endonuclease on DNA associated with each other, a nick remains to be sealed and give complete circular DNA. **What DNA ligase is doing?** It is an enzyme requires ATP or NAD^+ as a cofactor to catalyze ligation reaction. Ligase is processing ATP to generate AMP, and then AMP is making adduct with enzyme to form ligase-AMP complex. This complex is binding to the 3' and 5' of DNA bearing nick and bringing them together. AMP is released and phosphodieaster linkage is formed between 3' and 5' end to seal the nick (Figure 13.4).

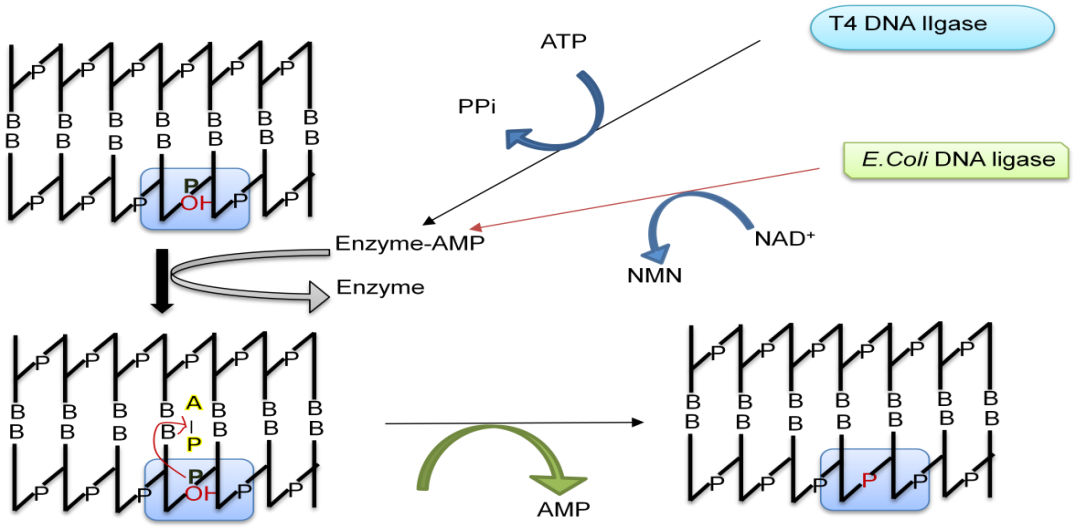

Figure 13.4 : Mechanism of DNA ligase.

3. Alkalike phosphatase- Digested linear plasmid containing cohesive ends on both sites with phosphate has a tendency to recircularize (Figure 13.5). Removing terminal phosphate group prevents this possibility and for this purpose, alkaline phosphatase is used. Alkalike phosphatase removes 5'-terminal phosphate groups and in this condition, only in the presence of insert DNA as it will supply phosphate group at both ends to facilitate the ligation reaction (Figure 13.5).

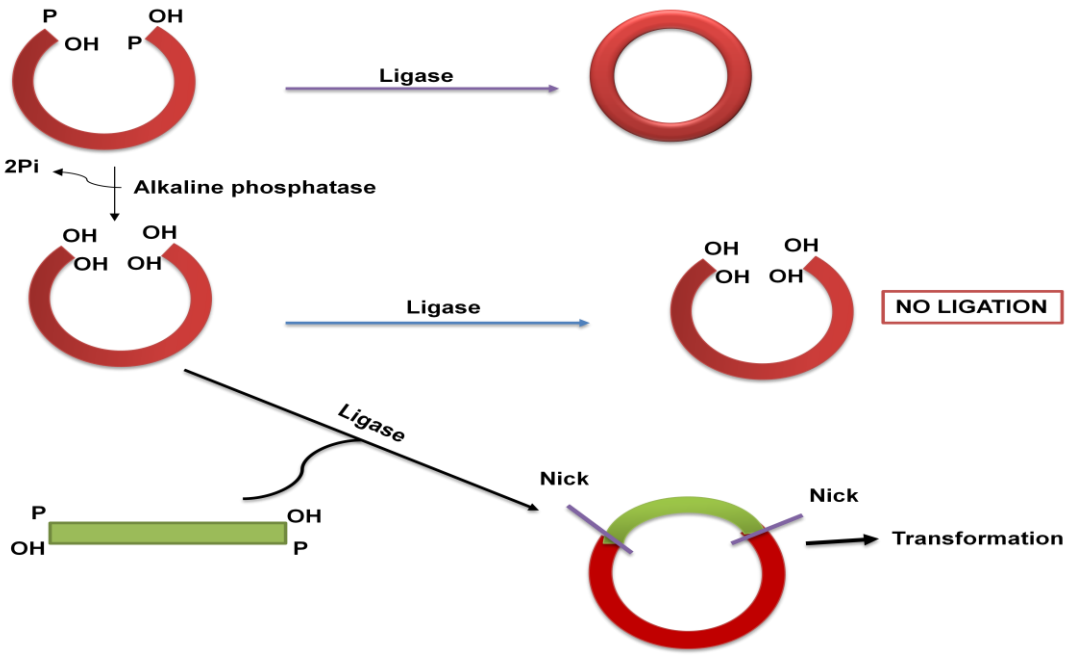

Figure 13.5 : Usage of Alkaline phosphatase in cloning.

Lecture 15: Polymerase Chain Reaction (Part-I)

Introduction : Polymerase chain reaction (PCR) is used to amplify a DNA sequence to produce millions of copies. Kary Mullis discovered the PCR and got Nobel Prize in Chemistry in 1993 for his discovery. Since then, PCR has been used in various applications in medicine, animal science, plant science, food science etc. The different events to develop present day PCR is given in Table 15.1.

TABLE 15.1: DIFFERENT EVENTS IN DEVELOPMENT OF PCR

Year	The gradual breakthrough from the discovery of DNA structure to the invention of modern PCR
1950	Discovery of mechanism of DNA Replication by Arthur Kornberg. He discovered the first DNA polymerases and other factors like helicase and primers.
1976	Isolation of thermostable DNA polymerase from *T.aquaticus*.
1983	Mullis synthesized DNA oligo probes for Sickle cell anemia mutation.
1983	Repeated thermal cycling was first used for small segment of cloned gene.
1984	Mullis and Tom White tried designed experiments to test PCR on genomic DNA but the amplified product was not visible in agarose gel.
1985	Patent was filed for PCR and its applications focusing on sickle cell anemia mutation.
1985	The use of thermostable DNA polymerase in PCR was started. Out of only two enzymes (Taq and Bst) known at that time, Taq was found more suitable for PCR.
1985	First announcement of PCR technique in Salt Lake City.
1985-1987	Development of instrument for PCR and its reagents.

Principle of the technique: The whole process of PCR involves three main events, Denaturation, Annealing and Elongation (Figure 15.1). A DNA fragment of interest is used as a template and a pair of primers which are short oligonucleotides complimentary to the both strands of the template DNA. The purpose of primer is to initiate the DNA synthesis in the direction of 5' to 3'. The number of amplified DNA or the amplicons increases exponentially per cycle thus one molecule of DNA gives rise to 2,4,8,16 and so forth. This continuous doubling is carried out by a specific enzyme called DNA polymerase which sits at the unfinished double stranded DNA created by template DNA and primer. For further extension of the DNA, the polymerase enzyme require supply of other DNA-building blocks such as the nucleotides consisting of four bases Adenine (A), Thymine (T), Cytosine (C) and Guanine (G). The template, primer, polymerase and four bases are the main components for polymerase chain reaction.

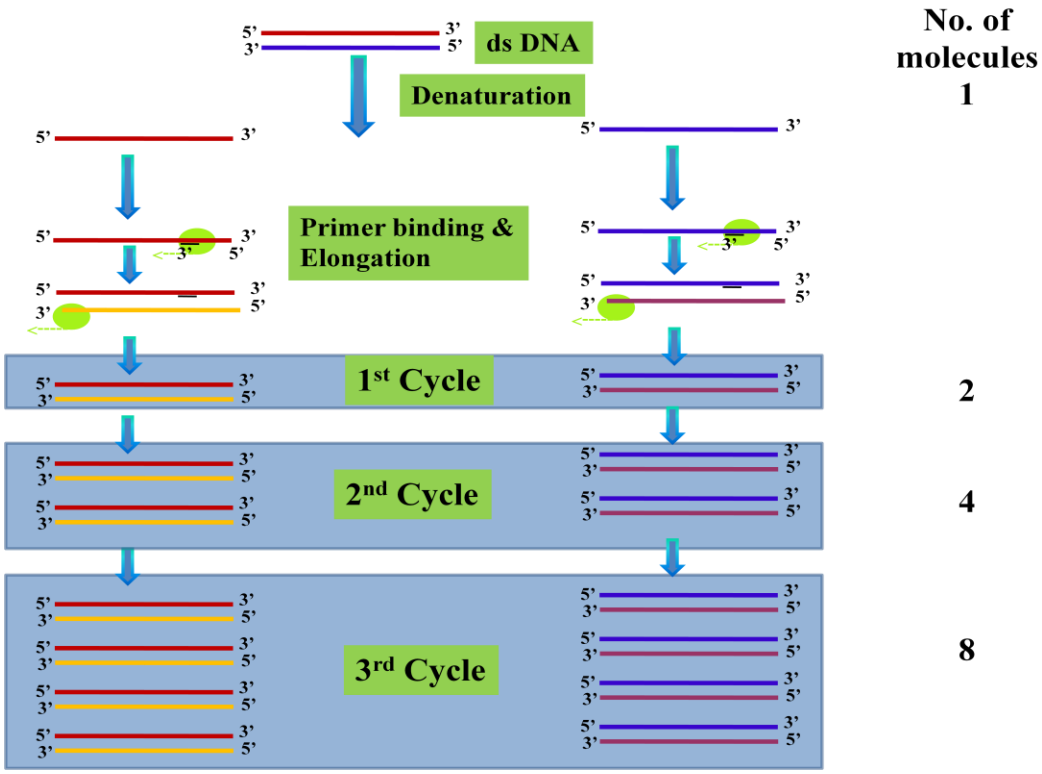

Figure 15.1: Basic Principle of polymerase chain reaction (PCR).

Methodology: PCR has three major events (Denaturation, Annealing and Elongation) to complete the amplification process (Figure 15.1). The complete process of PCR is as follows-

1. Initial denaturation: Heating the PCR mixture at 94°C to 96°C for 10min to ensure complete denaturation of template DNA. It is followed by the cyclic events which has different steps as described below:

A. Denaturation: This is the first step in which the double stranded DNA template is denatured to form two single strand by heating at 95°C for 15-30 secs.

B. Annealing: This is the annealing step where at lower temperature (usually 50-65^{0}C) primers are allowed to bind to template DNA, annealing time is 15-30 secs and it depends on the length and bases of the primers. Generally annealing temperature is about 3-5°C below the melting temperature (T_m) of the pair of the primers is used.

C. Elongation: This is the synthesis step where the polymerase perform synthesis of new strand in the 5' to 3' direction using primer and deoxyribonucleoside triphosphates (dNTPs). An average DNA polymerase adds about 1,000 bp/minute. Step 1,2,3 makes one cycle and in general 35-40 such cycles are performed in a typical PCR amplification.

2. After the cycles are completed, the reaction is held at 70-74°C for several minutes to allow final extension of the remaining DNA to be fully extended.

3. The reaction is complete and the resulting amplified nucleic acids are held at a low temperature (~4°C) until analysis.

Reagents: The reagents required for a complete PCR reaction volume is given in the table

Reagents	**Amount required**
Template DNA	1pg-1ng for viral or short templates 1ng-1µg for genomic DNA
Primers (forward and reverse primers)	0.1-0.5µM of each primer
Magnesium chloride	1.5-2.0 mM is optimal for *Taq* DNA polymerase
Deoxynucleotides (dNTPs)	Typical concentration is 200 µM of each dNTP
Taq DNA Polymerase	0.5–2.0 units per 50 µl reaction

Instrumentation: Thermal cycler is the instrument that carries out the amplification via polymerase chain reaction (Figure 15.2). Usually the three main events are repeated for 30-40 cycles to obtain detectable amount of product at the end of the cycles**.** The automated system performs the cyclic temperature changes required for enzymatic amplification of specific DNA segments in vitro using this PCR. The device has a thermal block with holes where tubes containing reaction mixtures can be inserted. The cycler varies the temperature of the block in discrete, pre-programmed steps using peltier effect.

Primers: A primer is a short oligonucleotide that serves as a starting point for DNA synthesis. In PCR, two primers are required to bind to each of the single stranded DNA (obtained after denaturation) flanking the target sequence. These are called Forward and Reverse primers. They primers are designed in such a way that they

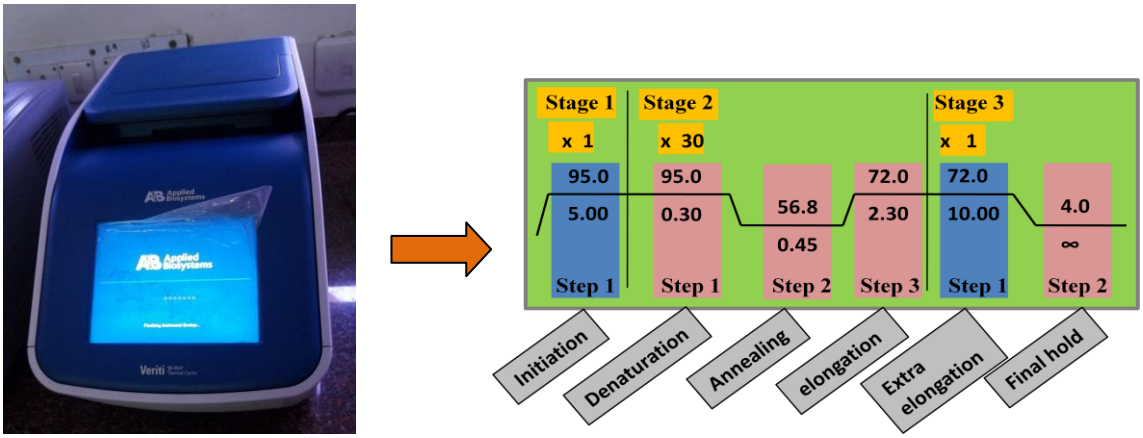

Figure 15.2. Representation of thermal cycler instrument showing the position of sample and schematic diagram of 30 cycle PCR.

have a sequence complimentary to the sequence in the template DNA. Two restriction enzymes sites are added at the 5' end of each of the primer to facilitate cloning. The chosen restriction enzymes will not cut DNA fragment (non-cutters). Typically 3 to 4 nucleotides are added at the end of the restriction sites to allow efficient cutting by restriction enzymes.

Primer Designing and criteria: For a specific amplifications in PCR, good primer design is essential. The following parameters needs to be considered while designing a primer:

1. Primer length: Oligonucleotides between 18-24 bases is the ideal length which is long enough for adequate specificity and short enough for primers to bind easily to the template at the annealing temperature.

2. Primer melting temperature (T_m): Primers with melting temperatures in the range of 52-58^oC generally gives the best results. The GC content of the sequence gives a fair indication of the primer $T_{m.}$ The two primers should be prepared in such a way that their T_m difference should not be more than 2°C otherwise it will result in poor annealing efficiency. T_m can be calculated by the following formula:

FOR A PRIMER LENGTH <14 NUCLEOTIDES	FOR A PRIMER LENGTH >13 NUCLEOTIDES
Tm = 4°C x (number of G's and C's in the primer) + 2°C x (number of A's and T's in the primer) **For salt adjusted Tm calculation,** **Tm = T_m = 81.5°C + 16.6°C x (log_{10}[SALT]) + 0.41°C x (%GC) – 675/N**	**T_m = 64.9°C + 41°C x (number of G's and C's in the primer – 16.4)/N Where, N is the number of nucleotides in the primer**

3. Primer annealing temperature (T_a): Too high T_a will produce insufficient primer-template hybridization resulting in low PCR product yield while too low T_a will lead to non-specific PCR products caused by a high number of base pair mismatches. Since T_a is the function of T_m so can be calculated with respect to melting temperature as given below:

T_a = 0.3 x T_m(primer) + 0.7 T_m (product) – 14.9 Where,
T_m(primer) = Melting Temperature of the primers,
T_m(product) = Melting temperature of the product

4. GC Content: The number of G's and C's in the primer as a percentage of the total bases should be 40-60%.

5. GC clamp: As GC forms a stronger bond than AT, the number of GC content at the 3' end of the primer should not be more than 3 otherwise it will result in non-specific tight binding at regions where G and C are abundant.

6. Primer secondary structures: Primer secondary structures arise as a result of intra or intermolecular attraction within the primer or with other primers which eventually reduce the yield of amplification as the availability of single stranded primers will be limited for PCR. The various types of primer secondary structures are as follows:

Hairpins: Hairpins are loop structures formed by intramolecular interaction within the primer. Optimally a 3' end hairpin with a ΔG of -2 kcal/mol and an internal hairpin with a ΔG of -3 kcal/mol is tolerated generally.

Dimers: A primer dimer is a structure forming a double-strand like structures which is formed by intermolecular interactions between the two primers. If interaction is formed between two homologous or same sense primers, it is called self-dimer whereas if interaction is formed between two different primers, it is called cross-dimer. Optimally, a 3' end self dimer with a ΔG of -5 kcal/mol and an internal self dimer with a ΔG of -6 kcal/mol is tolerated.

Repeats and runs: Repeats are consecutive occurance of di-nucleotide whereas runs are continuous stretch of

single nucleotide. A maximum number of repeats and runs accepted is 4 di-nucleotide and 4 base pairs respectively.

Primer-template homology: Primers should be designed in such a way that there should be no homology within the template other than the target site. This will result in non specific binding and amplification.

Analysis of PCR results: Once PCR cycle is complete, the amplified product is loaded in the agarose gel and observed after ethidium bromide staining under UV light source (Figure 15.3). A water blank reaction is included to monitor the cross contaminating DNA source as template. The percentage of agarose gel depends on the size of DNA to be visualized. Generally 0.8-1% agarose gel is used for analyzing 0.5-5 kb amplified DNA while a DNA of larger size or genomic DNA is visualized in gel as low as 0.5%.

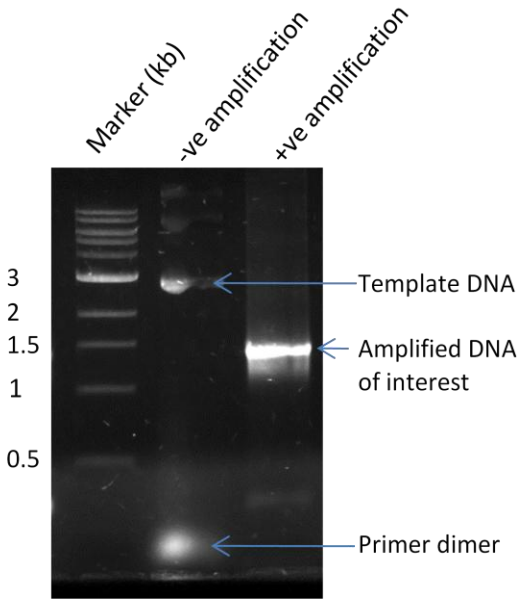

Figure 15.3: Analysis of PCR product on a agarose gel.

Lecture 16: Polymerase Chain Reaction (Part-II)

APPLICATIONS OF PCR: An over-view of the application of PCR in different fields is given in Figure 16.1.

PCR in human medicine: PCR technology has become an essential research and diagnostic tool for improving human health and quality of life. It allows the detection of infectious organisms just from one cell by amplifying specific region of the genetic material. Some important areas in medical research where PCR technology is employed include the following:

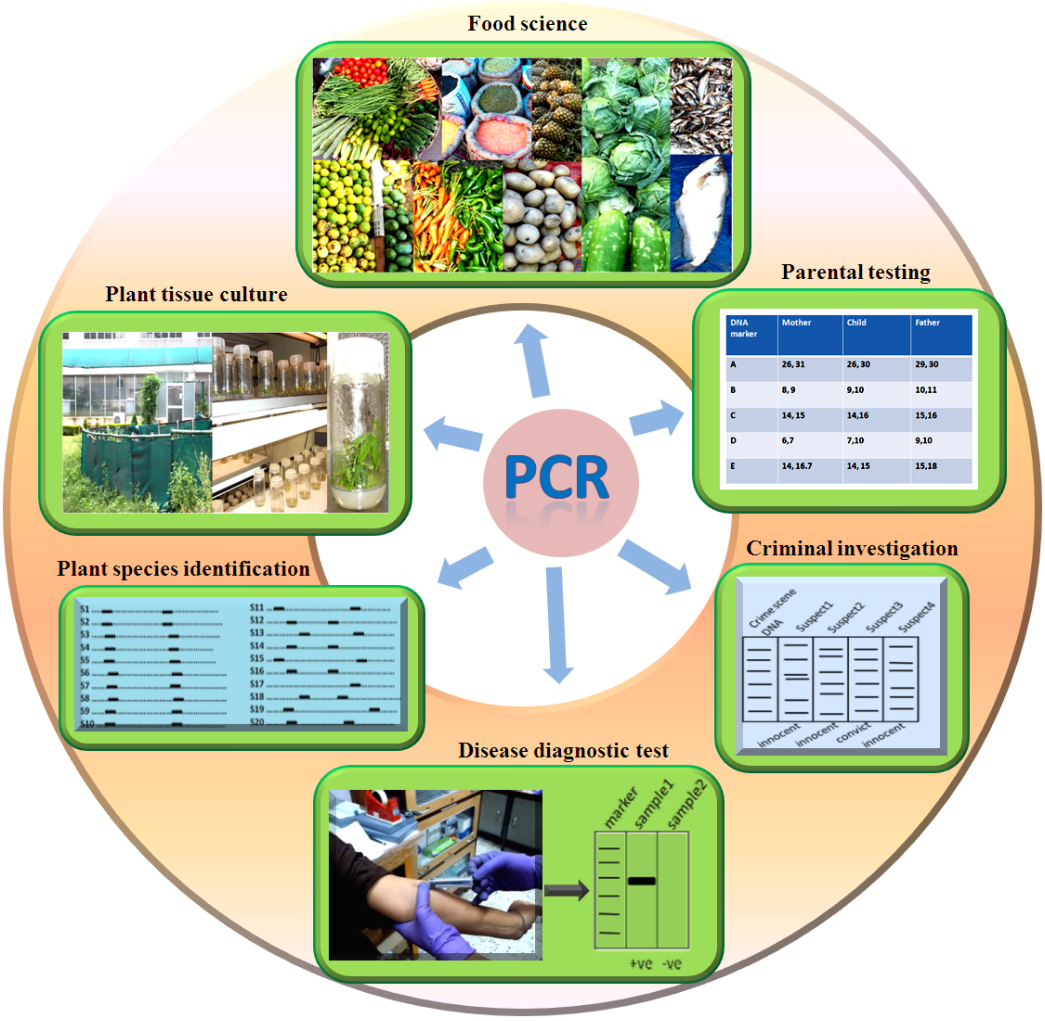

Figure 16.1: Applications of PCR in various fields of life.

PCR in infectious disease: PCR technology has become the basis for a broad spectrum of clinical diagnostic tests for various infectious agents, including viruses and bacteria (Figure 16.2). Besides detecting the presence of pathogens, PCR allows us to quantify the amount of pathogens present in patient's blood and this helps to monitor the progression of infection or response to drug treatment. PCR has enabled the development of diagnostic tests for many diseases such as, HIV-1, Hepatitis B and C viruses, Human Papillomavirus, *Chlamydia trachomatis*, *Neisseria gonorrhoeae*, *Cytomegalovirus*, *Mycobacterium tuberculosis*.

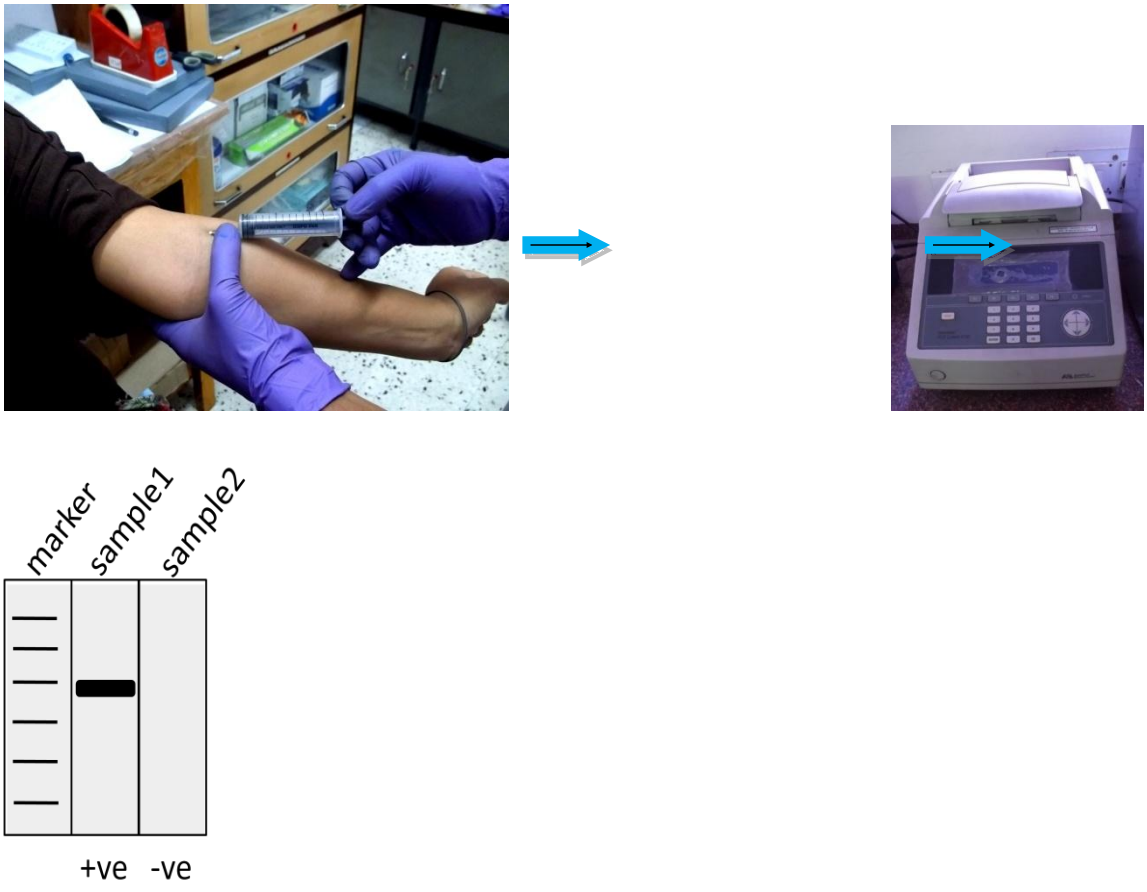

Figure 16.2: Use of PCR in HIV-1 test. The blood from patient is drawn and the viral DNA is amplied using PCR. The result is shown in gel. Amplification of specific fragment length indicates +ve result while no amplification means –ve result or absence of virus.

PCR and genetic testing: PCR technology has recently become a powerful tool to detect disease associated gene to predict the presence of heart disease and cancers. Knowledge of disease associated gene in the person predisposed to these disorders have a chance to control the problem much in advance.

PCR in plant science: There are various fields in plant science which requires the use of PCR technology for its accomplishment.

Plant species identification: PCR technique has also been employed in identification of plant species using species and group-specific primers targeting chloroplast DNA. These assays allowed identification of plants based on size-specific amplicons. plants belonging to the same family has close primer-binding sites and hence same amplicons size while plants belonging to different species and groups have different primer-binding sites and hence will result in different amplicons size.

PCR in tissue culture: It is used in analysis of DNA and specific genes in plant cells at different stages of re-generation during in vitro culture along with RAPD (random amplification of polymorphic DNA) technology. The level of polymorphism in regenerated plants could be revealed by these dual techniques.

PCR in veterinary parasitology: Owing to its rapidity and sensitivity as compared to antibody-based diagnosis, PCR met its uses in almost all aspects of biological work including veterinary clinical diagnosis. Some examples of PCR applications in veterinary parasitology:

Aujeszky's disease (pseudorabis) virus of pigs: This virus causes abortion and mortality in piglets. This disease has a latent period where there is no symptom of infection making it difficult to eradicate the disease completely. For this reason, PCR is considered to be appropriate tool for detecting latent cases of Aujeszky's disease.

Bovine leukemia virus (BLV): This virus causes enzootic bovine leukosis. PCR assay for detection of BLV was developed in 1991.

Bovine viral diarrhoea virus (BVDV): This virus is not only fatal to cattle but also causes contamination in calf serum used in cell culture work thus leading to contamination of vaccines and pharmaceutical products.

Besides the above examples, PCR has been used in routine diagnosis of veterinary virus such as Porcine parvo, Bovine papilloma type 1 and 2, Avian polyoma, Chicken anemia, Duck hepatitis, African swine fever, Channel catfish, Equine herpes type 1 and 4, Feline herpes, Alcelaphine herpes type1 etc.

PCR in Forensic applications: The most common use of PCR in forensic applications includes:

(1) Criminal investigation: A sample of blood, hair root or tissue left in the crime scene can be used to identify a person using PCR by comparing the DNA of the crime scene with that of suspect or with DNA database of earlier convicts (Figure 16.3). Evidence from decades-old crimes can be tested, confirming or defending the people originally convicted.

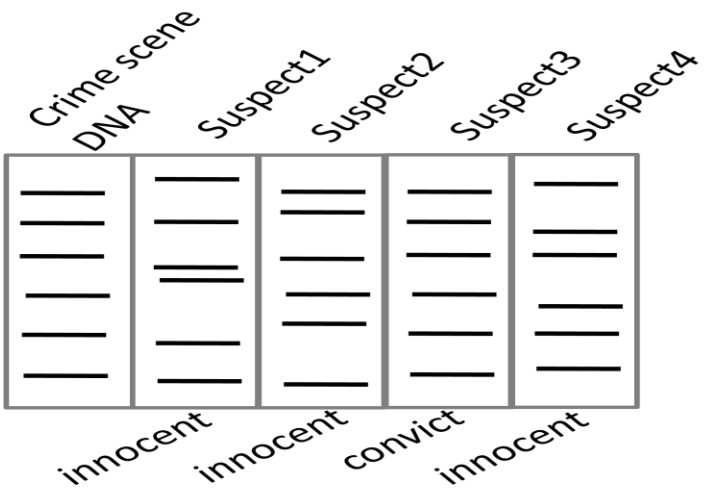

Figure 16.3: Use of PCR in criminal investigation. The DNA of suspects is amplified and the digested products are analyzed. The DNA from the crime scene matches that of suspect 3.

Parental testing: PCR technology is also used in finding the biological parents of adopted or kidnapped child where the DNA of a child is matched with close relatives (Figure 16.4). The actual biological father of a newborn can also be ruled out. In parental testing, short tandem repeats (STR) are used as markers where each person's DNA copies contain two copies of these markers one each from father and mother. These markers differ in length and sometimes sequence.

DNA marker	Mother	Child	Father
A	26, 31	26, 30	29, 30
B	8, 9	9,10	10,11
C	14, 15	14,16	15,16
D	6,7	7,10	9,10
E	14, 16.7	14, 15	15,18

Figure 16.4: Use of PCR in Parental testing is done by comparing the DNA markers (given as A→E for convenience) of mother, child and father. The child shares half of the DNA markers from each of the parents.

PCR in research applications: Biological research requires molecular biology techniques as its starting material and so forth which cannot be accomplished without the use of PCR.

DNA cloning: PCR helps to amplify specific DNA from a genome and the amplified DNA can be inserted into a vector for transformation and expression. These inserts can further be confirmed by PCR method.

DNA sequencing: PCR assists the task of DNA sequencing from patients with genetic disease mutation.

Sequence-tagged sites: This is a process where PCR is used as an indicator if a particular segment of a gene or genome is present in a particular clone. This application is vital in mapping the cosmid clones to be sequenced by the Human Genome Project.

Phylogeny analysis: The phylogeny of organisms like plants, animals and other lower organisms can be traced out by DNA analysis. The origin of unknown samples like the recovered bones of early men can also be rules out.

Quiz

Q1: What is the advantage of PCR as diagnostic tool with notable example?

Q 2: What are the enzymes used besides Taq DNA polymerase in PCR?

Q3: Describe how primer with low Tm and secondary structure affects the PCR yield?

Q4. Describe the different applications of PCR in drug discovery?

Q5. What is the role of Mg^{2+} in PCR reaction ?

Lecture 17: Cloning Vectors-I

Introduction: In the steps of cloning discussed in earlier lecture, a vehicle DNA is required to carry foreign DNA to generate a recombinant construct so that it should allow easy amplification of chimeric DNA in host. Vehicle DNA used for this purpose in genetic engineering is known as **'vector'**. In today's lecture we will discuss the properties of different vector used for host strains.

Criteria of a good vector : The vector DNA has a two main responsibility: (1) ability to carry foreign DNA, (2) ability to replicate in the host. Hence to fulfil these responsibility, a number of properties are desirable. Few crucial properties are as follows-

1. Low molecular weight-The low molecular weight or size confers a number of advantages. (1) small size vector is robust towards shear stress and easier to handle. In addition, after ligating foreign DNA into the vector, the size of the resulting recombinant DNA will be small and it will be easier to deliver the recombinant DNA into the host cell.

2. Post entry into the host should give phenyotypic changes-Another important feature is that vector DNA should give additional phenotypic changes in the host cell so that recognition of transformed cells will be easier.

3. Muliple cloning site with unique restriction site- A short stretch of DNA on vector DNA containing restriction site for possible site for insertion of foreign DNA is desirable.

4. High copy number-A high copy number is desirable as it gives high amount of DNA after growing host cells.

Different vectors: As vector needs to replicate in different host strain, vector needs special additional structural features to make it suitable for a particular host strain. **Why one vector doesn't rplicate in different host strains?** Replication of vector DNA is controlled by the orgin of replication and it need to be recognized by the host factor especially DNA polymerase to perform replication. Consequently, there are different types of vector DNA to suits the cloning of a foreign DNA in a particular host strain.

The Different host specific vectors, we are going to discuss as follows-

Bacterial Plasmid

Phage based vectors

Yeast vectors

Mammalian vectors

Bacterial Plasmid : Plasmid are widely been used for cloning of foreign DNA into the bacteria as host strain. Before getting into the details of discussing bacterial plasmid we will discuss the basic properties of plasmids.

Different forms of plasmids: Bacterial plasmid is a double stranded circular DNA exists in 3 different forms (Figure 17.1). If the both strands of circular double strands are intact then it is called as **covalently closed circles** (CCC) where as if one of the strand has nick, then it acquire the conformation of **open circle** DNA (OC, DNA). During the isolation of plasmid DNA from bacteria, covalently closed circular DNA losses few number of turns and as a result it acquire **supercoiled** configuration. The interchange between these different forms are possible under the in-vitro or in-vivo conditions, such as DNA gyrase produces additional turn into the circular DNA to adopt supercoiled conformation.

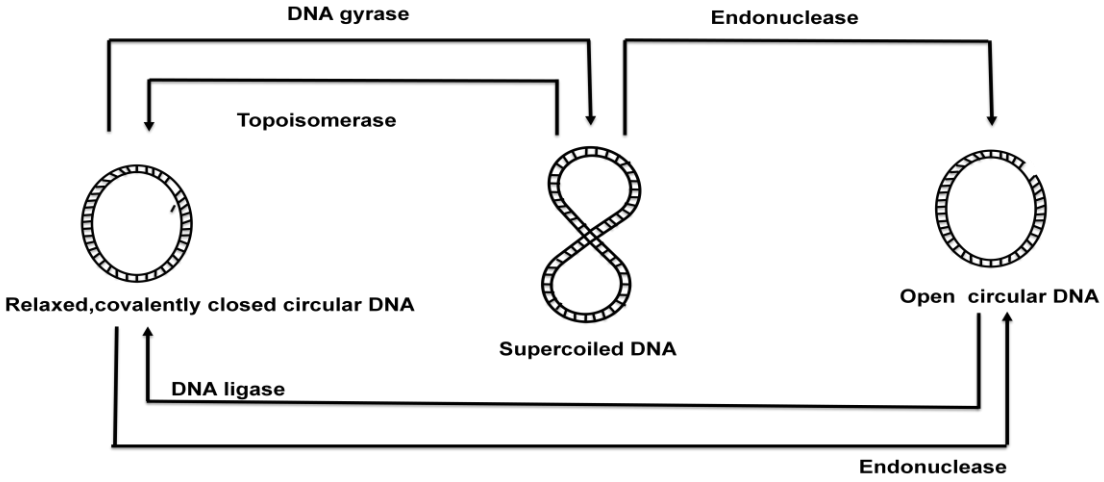

Figure 17.1: Different forms of plasmids

Features of different plasmids: There are minimum molecular components to assemble bacterial plasmid to perform the function of vector are as follows-

1. Origin of replication-Like any other replicating DNA, plasmid DNA needs its own independent origin of replication to provide replication start site to make more copies. It decides the range of bacterial host strain can be use with the particular plasmid vector. The plasmids containing ori region from Col E1 can be able to grow in limited bacterial species such as E.Coli etc. In contrast, plasmid containing ori from RP4 or RSF1010 can be able to grow in gram (-) bacteria and gram (+) bacteria.

2. Selection marker- Selection marker in the form of either antibiotic resistance gene or enzymatic gene is essential to give phenotypic changes in host after entry of the plasmid.

3. Promoter- Plasmid replication in host is performed by the host provided proteins such as DNA gyrase, helicase, polymerase and DNA ligase. But proteins required for conferring antibiotic resistance or enzyme use for selecting transformed host cells is present on plasmid and a promoter adjacent is required to express genes present on plasmid DNA. In addition, promoter is also needed to express gene present on foreign DNA.

pBR322: In early days, natural occurring plasmids such as col E1 and RSF1010 were used for cloning. But these plasmids had several limitation such as number of selection marker and recognition sites for restriction sites. To facilitate the cloning, a much improved cloning vector was produced and named as pBR322. It is produced after taking structural elements from RSF2124, pSC101 and pMB1. Plasmid pBR322 received ampicillin (Ap^R) and teracyclic (Tc^R) resistance gene from RSF2124 and pSC101. Origin of replication is derived from pMB1 and a detail of construction of pBR322 with multiple steps of recombination is given in the following article. [**Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene. 1977;2(2):95-113. PUBMED ID: 344137**]. A vector map of pBR322 with regions derived from different plasmid is given in Figure 17.2. pBR322 is a 4359 bp long plasmid and has 40 unique restriction sites (Figure 17.3). Eleven restriction sites are present within Tetracycline resistance gene and six sites are within ampicillin resistant gene. In addition, two sites are present within promoter of the Tetracycline resistance gene. Cloning any

DNA fragment into these sites will disrupt the resistance gene and as a result it can be used as a criteria for selecting recombinant plasmid. The details of different selection methods are discussed in future lecture.

Application of pBR322-

1. it is the most popular plasmid for cloning purpose.

2. it is used to study transcription and translation of prokaryotic gene.

3. it is the primary sources to design and construct improved plasmids for specific applications.

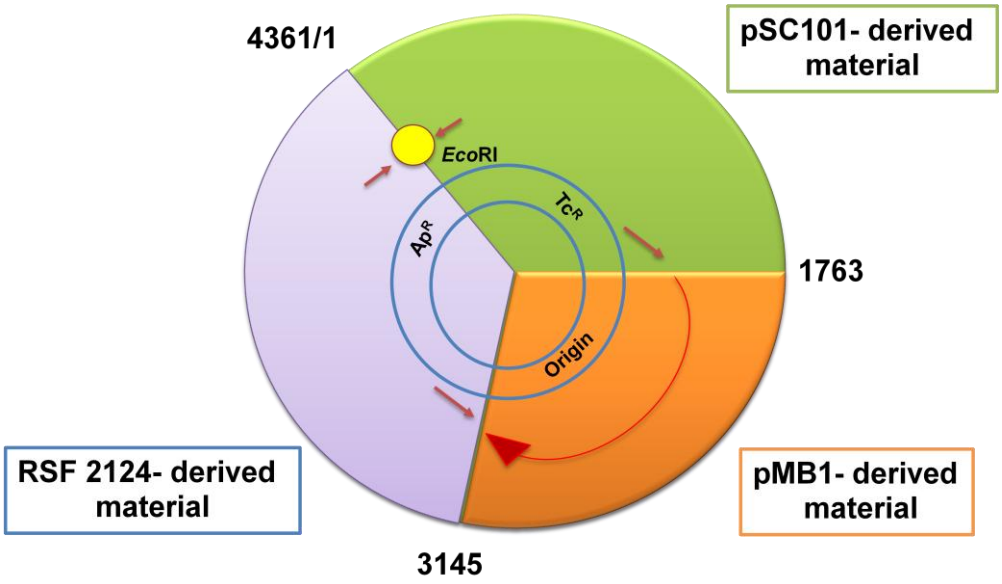

Figure 17.2: Construction of pBR 322 using structural elements from other plasmids.

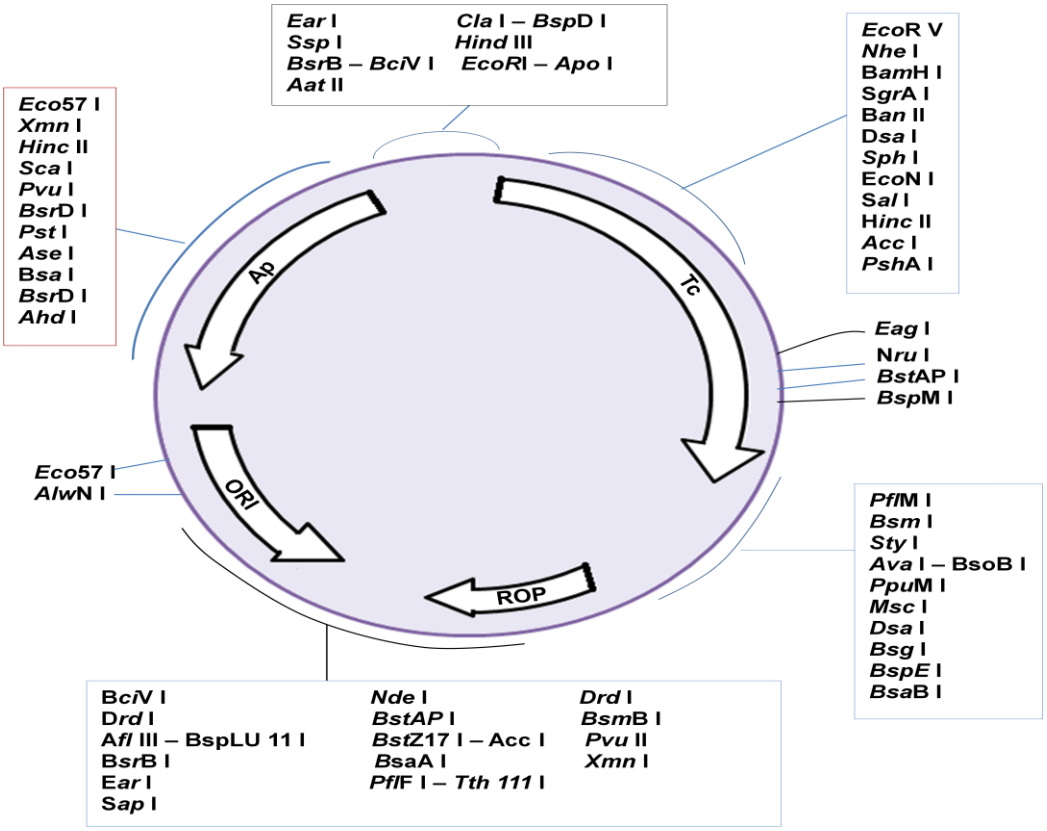

Figure 17.3: Vector map of pBR 322

pUC19: pBR322 was constructed in 1977 and since then large number of plasmids were derived from this plasmid to achieve specific requirement of an application. It was mainly been done by adopting different screening strategies than pBR322 and addition of other structural features. In the similar advancement, a multiple cloning sites were added to facilitate easy cloning without disrupting antibiotic resistance gene. pUC19 is the intial example of bacterial plasmid of small size (2.8Kb) containing multiple cloning site (Figure 17.4). The usual place to keep the MCS is always between initiation codon (AUG) and the codon 7. An MCS allow design of many cloning strategies as the large number of enzyme available for cloning. In addition, two enzymes from MS can be used to insert the foreign DNA without disturbing plasmid sequences. pUC 19 vectors also has a small stretch of DNA which encodes for rapid detection of an insert by blue-white screening. The details of different screening methods are discussed in future lecture.

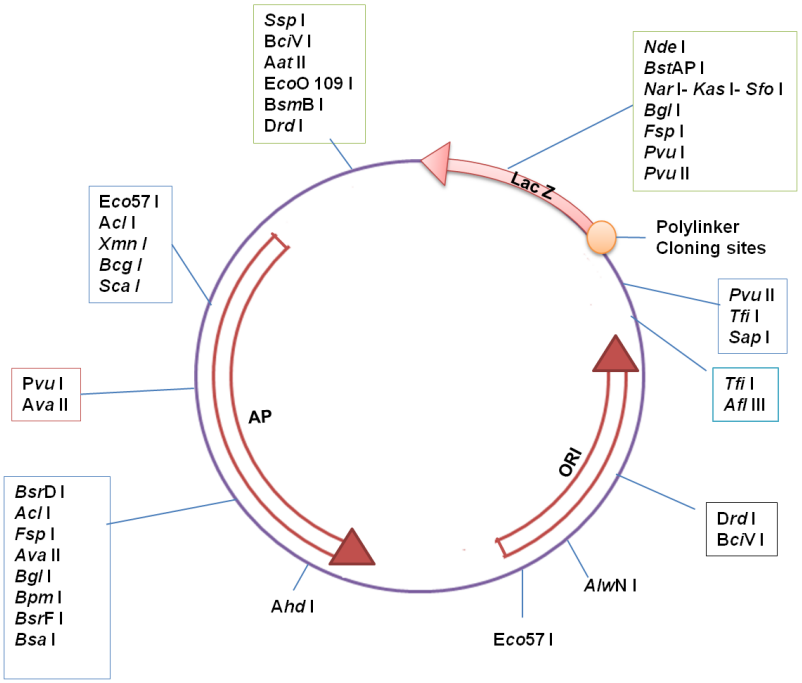

Figure 17.4: Vector map of pUC19

Isolation of plasmid from bacteria: The different steps in isolation in bacterial plasmid is given in Figure 17.5.

STEP 1: The bacteria containing plasmid was grown in suitable culture media in high density (~0.8 optical density). Each Bacterial cell contains chromosomal DNA, plasmid DNA and cellular proteins. The bacterial culture is collected by centrifugation at the bottom and resuspended in the solution I containing 50 mM glucose, 25 mM TrisHCl pH 8.0, 10 mM EDTA pH 8.0.

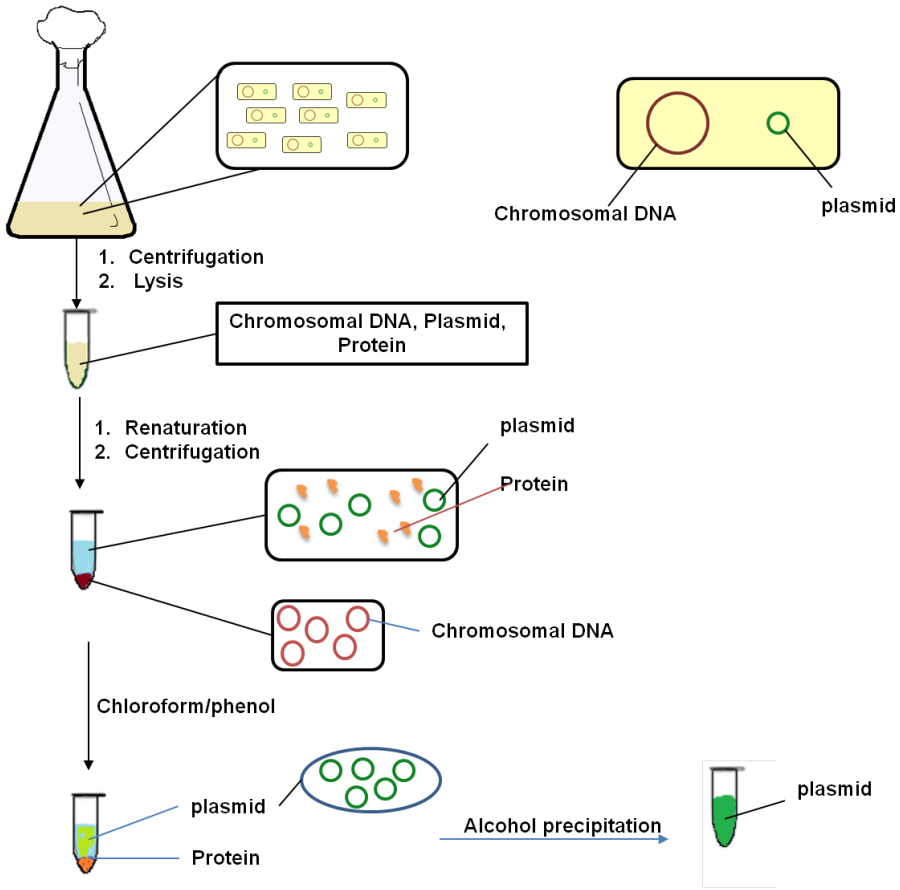

Figure 17.5: Steps in Plasmid Purification from bacterial culture.

STEP 2: Alkaline Lysis: Bacterial cells are treated with lysis solution II containing 0.2 N NaOH and 1% SDS. to lyse the cells and denature DNA (both chromosomal and plasmid DNA).

STEP 3: Renaturation: In 3rd step, denatured DNA is renatured with solution III containing potassium acetate, glacial acetic acid. In this step small DNA (plasmid) renature back quickly whereas chromosomal DNA remained denatured.

STEP 4: Deproteination: Resulting supernatant containing plasmid DNA and protein is treated with phenol: chloroform: isoamyl alchol mixture to remove protein in the precipitate where as plasmid remained in solution.

STEP 5: Precipitation : plasmid is precipitated by 100% alchol from the solution.

Lecture 18: Cloning Vectors-II

Bacteriophage λ based vector-Bacteriophage are the virus using bacteria as their host for replication. Bacteriophage λ is the virus of E.coli and have been used to develop vector for genetic engineering. Before discussing properties of bacteriophage λ based vector, it is important to understand bacteriophage biology.

Bacteriophage λ genome-Phage genome is a linear double stranded DNA of 48.5kb (Figure 18.1). On both end of the genome, it has a stretch of 12 nucleotides which are complementary to each other. These sites are known as "cos sites" and it allows circulation of virus genome after entering into the host cell. Genes are arranged in between these two cohesive ends and codes for proteins responsible for making head, tail, factor for recombination and process of lysogeny. The central region of the genome is non-essential and can be replace without much affecting growth and infectivity of the virus. As a result this region can be exploited to develop cloning vector with multiple approaches as discussed later in the lecture.

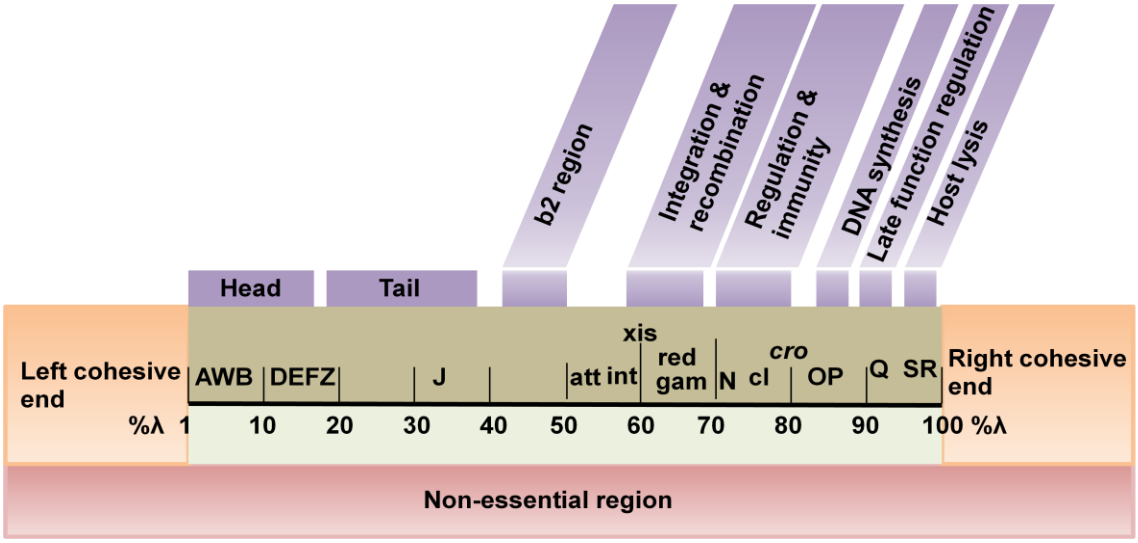

Figure 18.1: Map of bacteriophage λ chromosome.

Packaging of genome in virus particle-Virus genome is packed in a nucleoprotein complex to give virus particle. These virus particles bind to the host cells and deliver the genome for further replication. The steps in packaging are given in Figure 18.2. Phage genome is replicated by a rolling circle model to produce long genome where cos sites are present on regular interval. The 2 flanking cos sites and the DNA in between them constitute one virus genome or **monomeric unit**. In the presence of head precursor, the long genome is cleaved into the monomeric unit and encapsulated. Nicks are introduced on both strand of the genome to generate linear strand to serve as cohesive site to facilitate circularization in host cell.

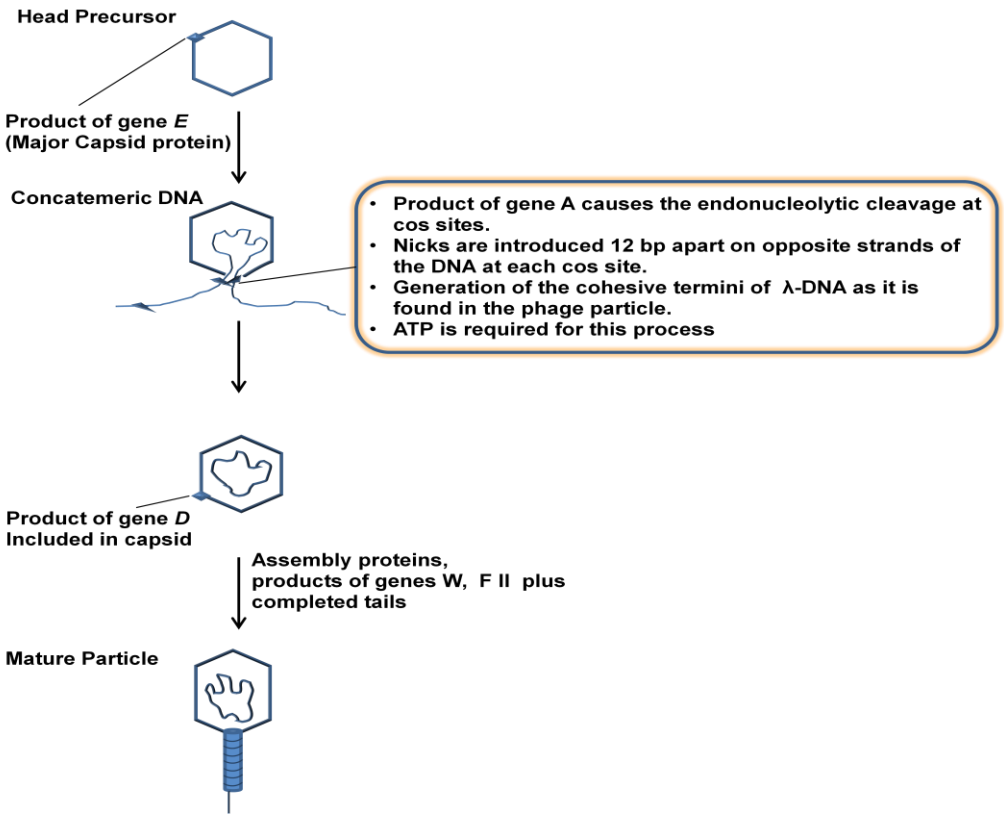

Figure 18.2: Packaging Mechanism of bacteriophage λ

bacteriophage λ cloning vector: The bacteriophage λ cloning vector has a middle segment responsible for insertion/excision (I/E Region) and this region can be replace with the foreign DNA with the help of two BamHI site present on the either side of I/E region. Hence, in a cloning strategy described in Figure 18.3, foreign DNA is put into the vector and then allowed to infect the bacteria. In the presence of I/E region, phage will integrate into the bacterial chromosome and continue lysogeny cycle. But when I/E region is disrupted or replaced with the foreign DNA, it will continue lytic cycle and form plaque. The examples of bacteriophage λ based vector are EMBL3, EMBL4 and their vector map is shown in Figure 18.4.

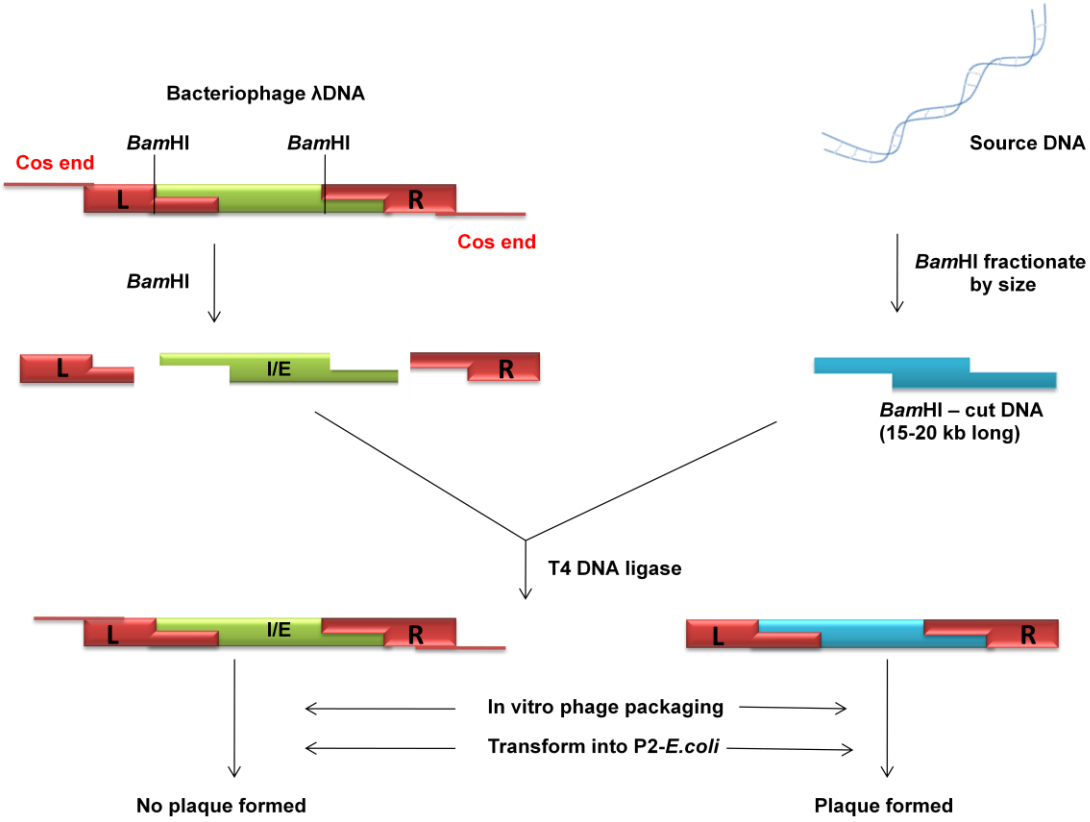

Figure 18.3: bacteriophage λ cloning system.

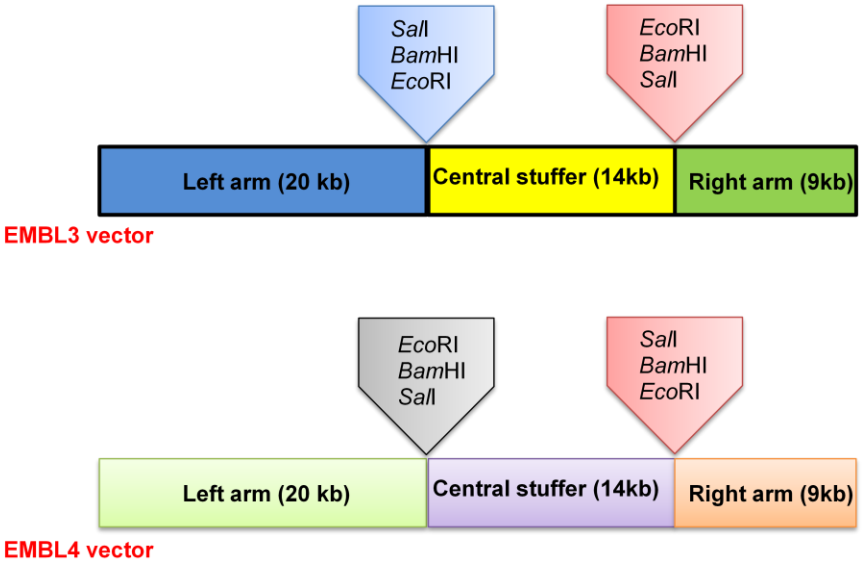

Figure 18.4 : Vector map of bacteriophage λ based vector EMBL3 and EMBL4.

Cosmids: Cosmids vector offer additional advantage over bacteriophage λ based cloning vector as they have bacterial origin of replication. They are chimeric cloning vector consists of region from a bacterial plasmid and bacteriophage λ. They contains flanking cos site and hence after their delivery into the host cell, it adopts circular form. As cosmids contains bacterial origin of replication, it can be maintained in bacteria as such. In addition, it has antibiotic resistance gene (tetracycline) and allow selection of transformed host cells. The cloning strategy follows the similar mechanism as discussed before for bacteriophage λ based vector and it is outlined in Figure 18.5. The example of cosmid vector is pLFR-5.

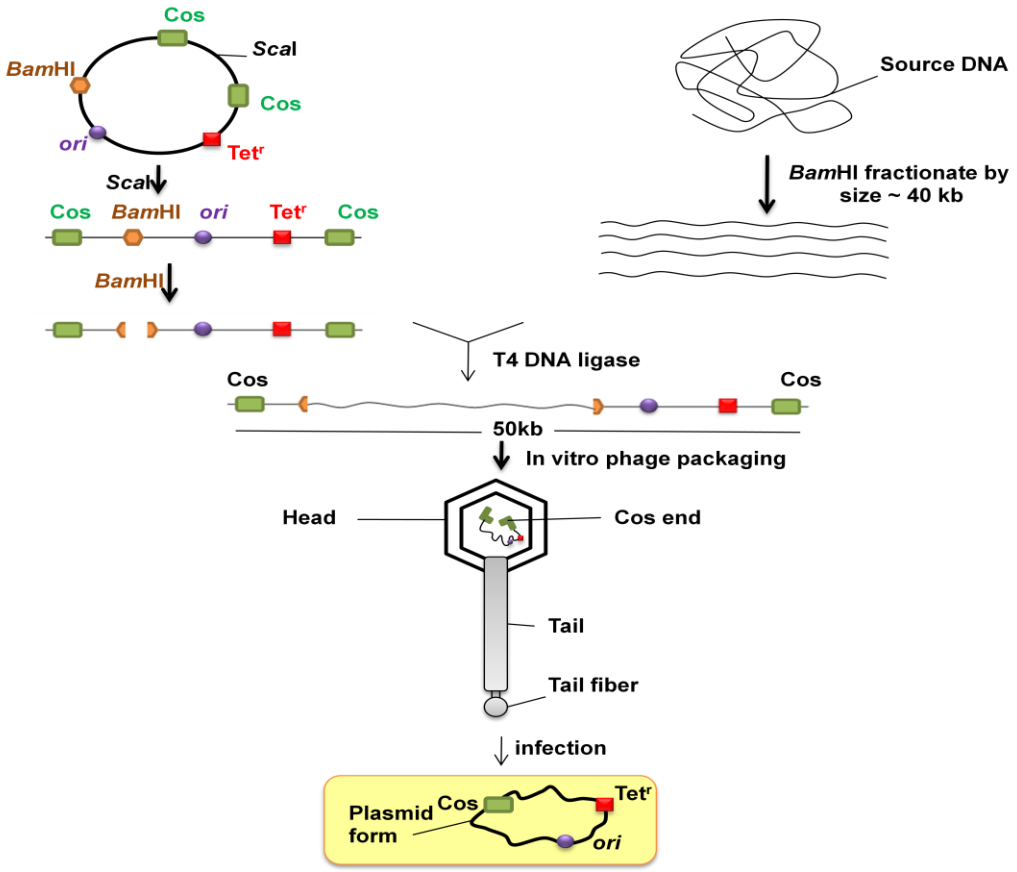

Figure 18.4 : Cosmid cloning system.

Lecture 19: Cloning Vector-III

Introduction: Prokaryotic vectors are good to express the proteins of eukaryotic or prokaryotic origin but in few specific cases, they are not well suited. Such as, eukaryotic protein is unstable, require special environment for folding or losses its the biological activity. In few cases especially in the case of production of therapeutic proteins, cross contamination of bacterial products may cause clinical problems in humans. Eukaryotic vector system is designed to clone and express gene in eukaryotic cells such as yeast (saccharomyces cerevisiae), insect and mammalian cell lines. There are two different types of eukaryotic vectors-

1. Vector as extrachromosomal DNA- These vector remains in eukaryotic cell as extrachromosomal DNA and express the protein.

2. Integration Vector- These vectors carry an integration site to facilitate recombination medited integration into the chromosomal DNA of the host cell.

In general, eukaryotic vector contains origin of replication from bacteria and eukaryotic cells. In addition, they also contain different selectable markers for prokaryotic and eukaryotic cells. These modifications allow to use and perform easy cloning procedure in bacterial host system to generate recombinant construct containing foreign DNA in vector. The basic features required for a vector discussed previously for prokaryotic system is also required for eukaryotic vector as well.

***Saccharomyces cerevisiae* vector system-**There are 3 types of yeast vector system. These all have couple of similar features such as presence of MCS, shuttle vector (origin of replication for E.coli and Yeast) and presence of selectable marker.

1. Episomal vector- Yeast episomal vector are constructed by combining bacterial plasmid either with yeast 2µ origin of replication or with autonomous replication sequence (ARS). The yeast vector containing ARS are high copy number plasmid but they are unstable in the absence of selection pressure. This can be over-come by adding centromeric sequence (CEN) but it affects the plasmid copy number and as a result it become a low copy number plasmid. ARS based yeast plasmids are not very popular for expression of protein. In contrast, 2µ based vector are very popular for heterologous protein expression. A representative 2µ based episomal yeast vector is shown in Figure 19.1. It is a 6.3kb plasmid with a copy number in the range of 50-100 per cell. These plasmids are much more stable than ARS based plasmids.

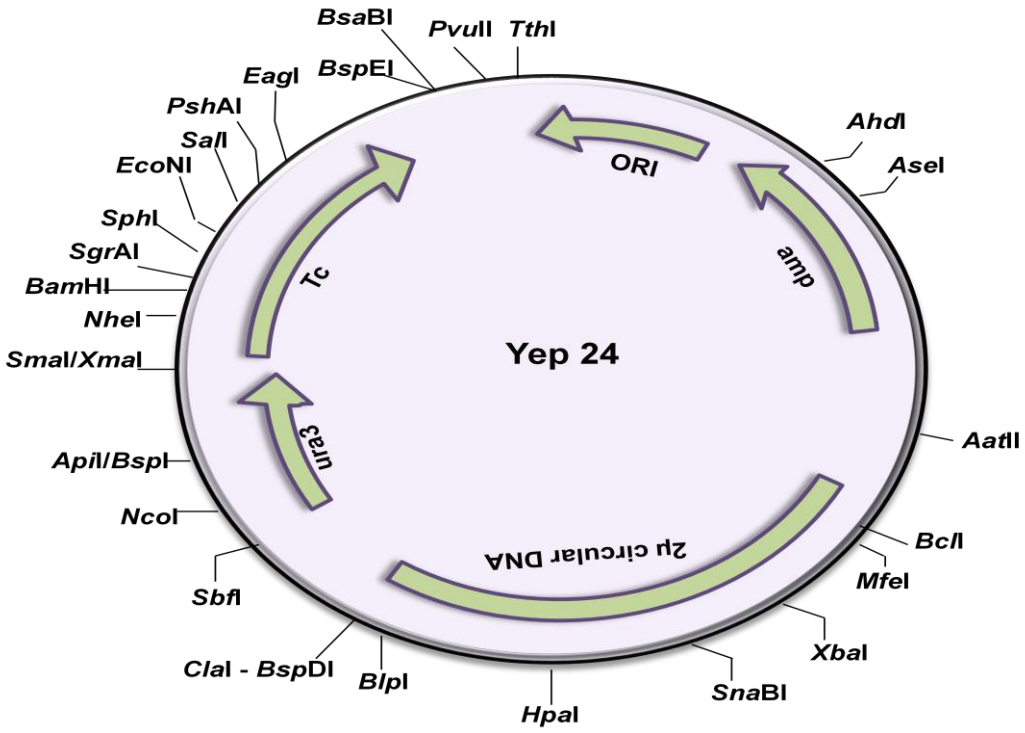

Figure 19.1: Vector map of episomal yeast plasmid Yep 24.

2. Integrating vector- Episomal yeast vectors are present as extra-chromosomal DNA and are unstable. This can over-come by integration of vector into the host chromosome. In yeast integration occurs by homologous recombination. In yeast integration plasmids contains target sequence for integration into chromosomal DNA, selection marker and bacterial origin of replication. Before vector delivery to the yeast, it is digested with the unique restriction endonuclease to produce linear DNA to increase the transformation efficiency and integration. In most of the cases integration is done in such a way that yeast chromosomal DNA remained intact and integration may not affect yeast growth. But in an alternate approach, a portion of yeast chromosomal DNA is replaced with the vector DNA through homologous recombination. These vectors are known as '**transplantment integration vector**'and they have foreign DNA, selection marker and homologous DNA to the region of chromosomal DNA to be replaced.

3. Yeast artificial chromosome- Yeast artificial chromosome (YAC) is the vector of choice used to clone very large DNA fragment (~100kb) to prepare genomic library. YAC vector is like a chromosome as it has ARS sequences, centromere sequence and telomere at the two ends to give stability. A typical YAC plasmid (pYAC) is shown in Figure 19.2. It has an ampicillin resistance gene (Amp^r) for selection in *e.coli* and an *e.coli* origin of replication for propagation in bacteria. In addition, it has ARS for replication, CEN for centromere function, and URA3, TRP1 for selection in yeast. URA3 and TRP1 is the crucial gene of uracil biosynthesis and tryptophan biosynthesis pathway. For cloning, YAC is digested with SmaI/BamHI, alkaline phosphatase to generate a linear plasmid DNA, now foreign DNA is added for ligation. The recombinant DNA will allow a yeast (Ura^-/Trp^-) to grow on uracil and tryptophan deficient media.

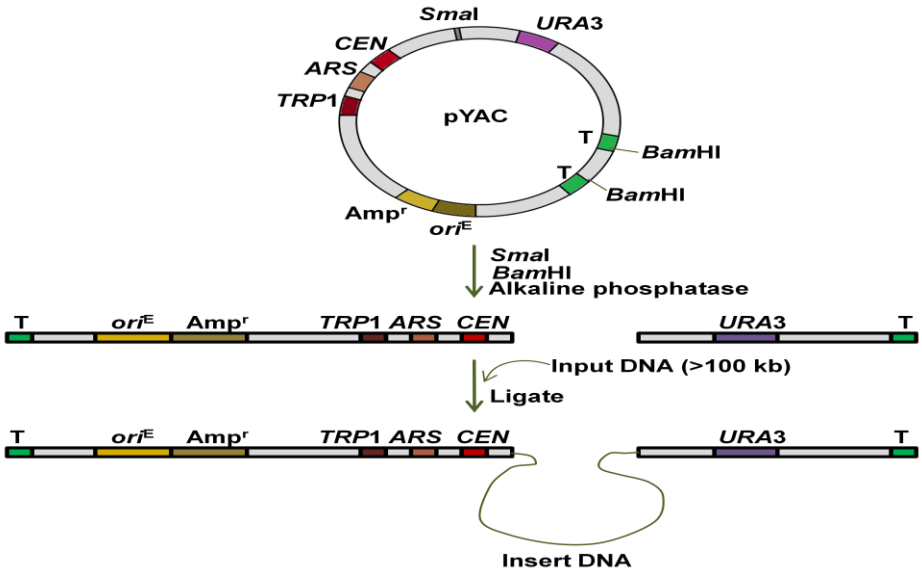

Figure 19.2: Vector map of YAC plasmid (pYAC) and YAC cloning system.

Baculovirus Vector-Baculovirus is a rod shape virus infecting invertebrate including insect cells. Post infection, virus is either released as free virions or many virus particles are trapped in a protein complex known as polyhedron. The protein responsible for trapping virus into polyhedron is polyhydrin and it help in transmission of virus from one host to other. The polyhydrin is not important for virus propagation but it is under very strong promoter to produce the protein in large quantities. Realizing this fact, replacement of polyhydrin gene with a foreign DNA fragment will allow expression of protein in large quantities. The baculovirus *Autographa californica* multiple nuclear polyhedrosis virus (AcMNPV) is used as a vector to express protein. The transfer vector map of AcMNPV is given in Figure 19.3. The gene of interest will be inserted into the cloning site placed adjacent to the promoter. It has polyhedron termination sequence down-stream to the cloning site to stop transcription of cloned gene. A more details of over-expression strategies will be discussed in future lecture.

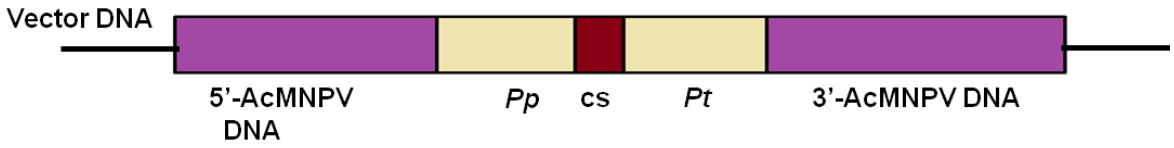

Figure 19.3: Structural elements of a baculovirus transfer vector.

Mammalian Vector- large numbers of excellent mammalian vectors are in circulation to clone eukaryotic gene for protein synthesis and study the transcription mechanism. A generalized scheme with the structural elements required to design mammalian vector is given in Figure 19.4. As discussed earlier, it contains a eukaryotic replication of origin from an animal virus such as SV40 from simian virus 40. A promoter to drive the expression of foreign gene and selection marker, other eukaryotic features such as polyadenylation, transcription termination site etc.

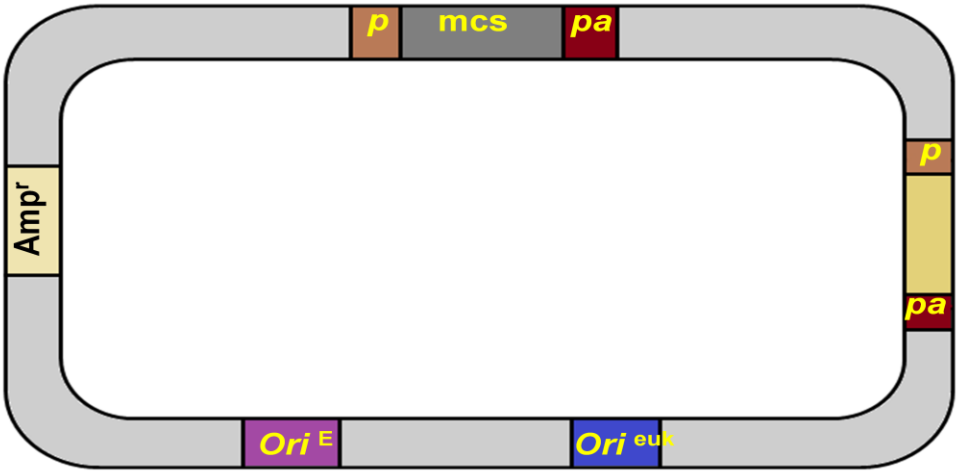

Figure 19.4: Structural elements of a mammalian expression vector.