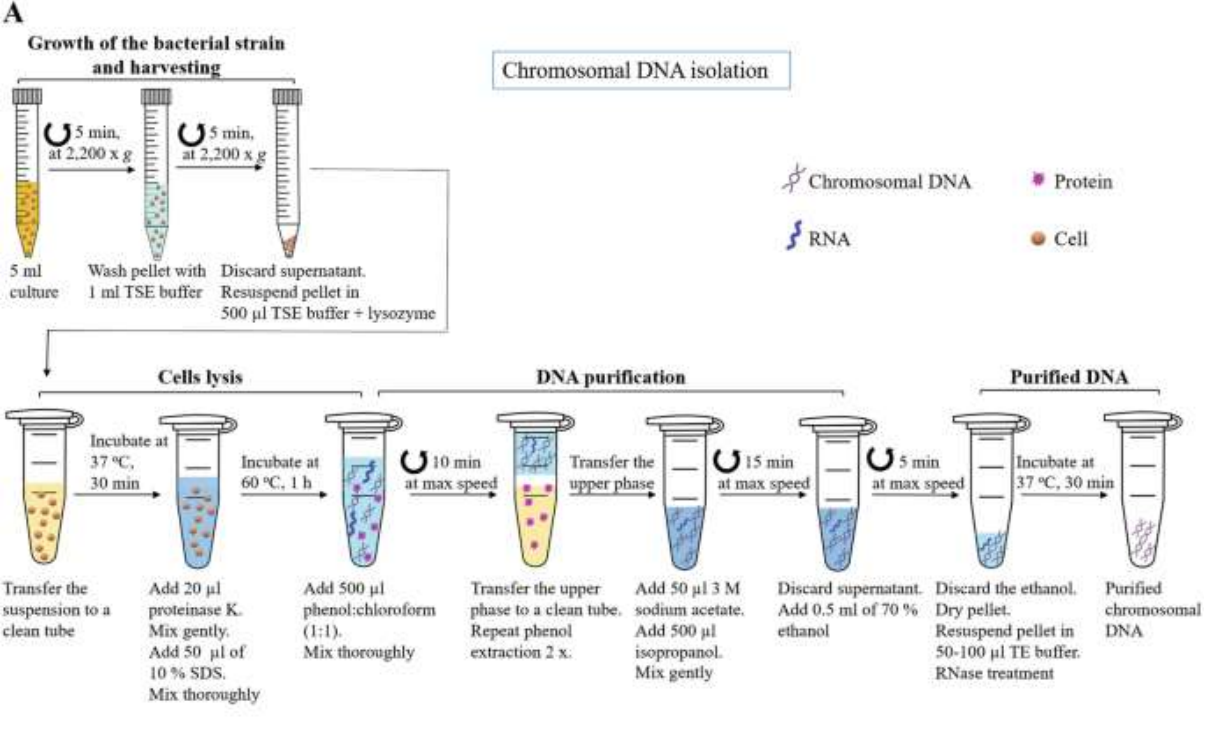
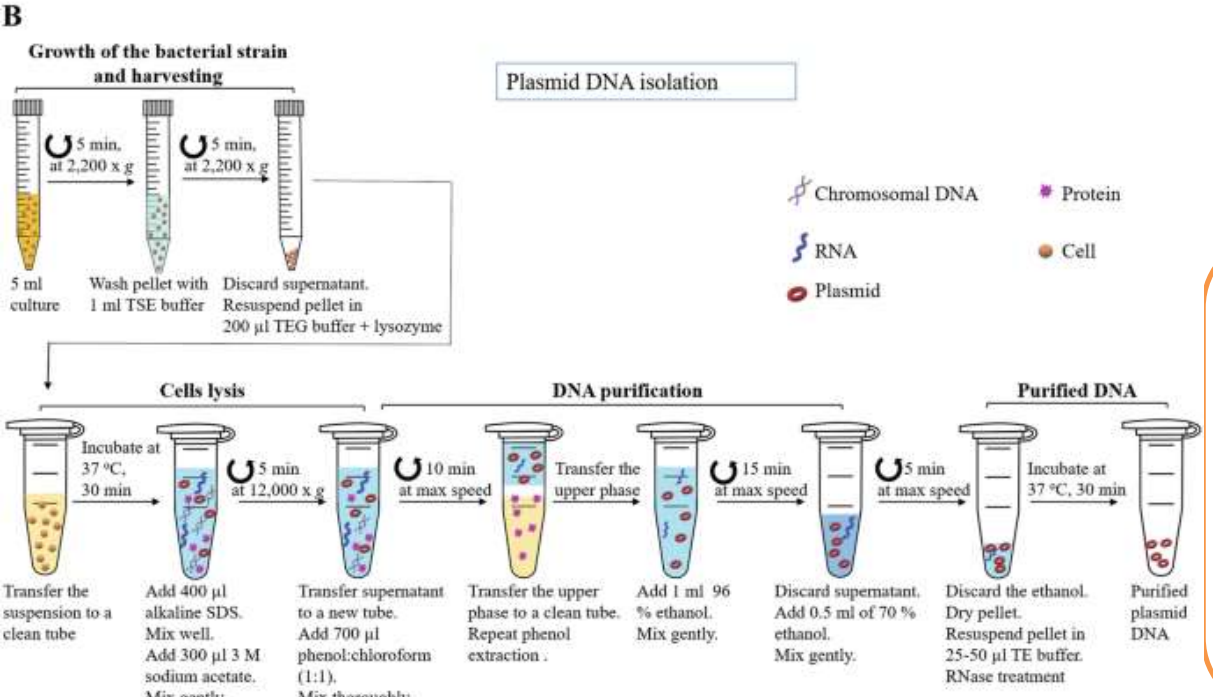
RECOMBINANT DNA TECHNOLOGY

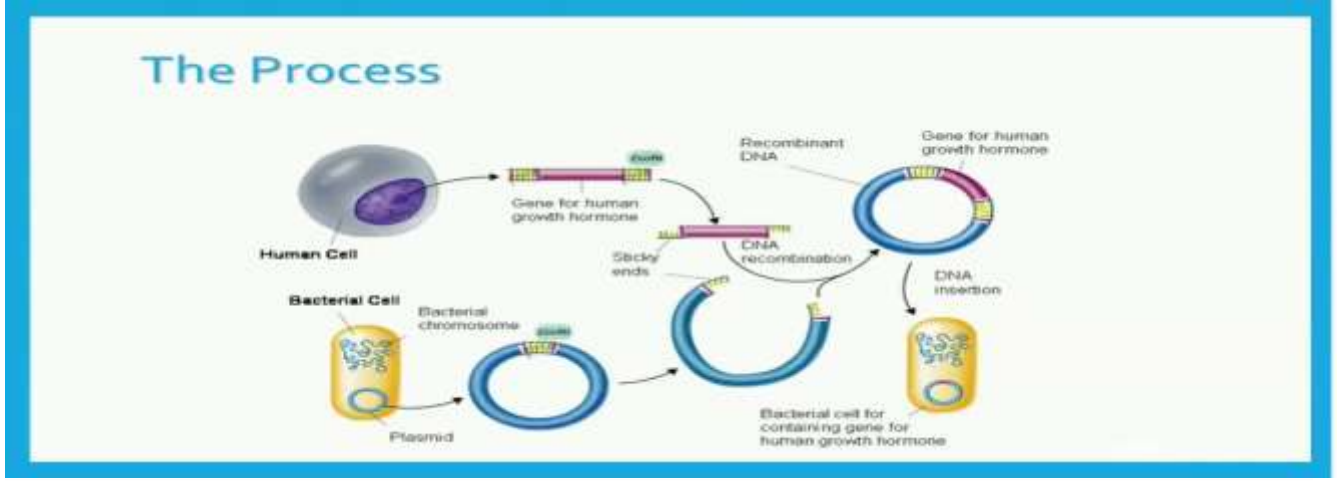
II M.Sc., MICROBIOLOGY
SEMESTER III

UNIT II (CLONING STRATEGIES)

SURESH S S RAJA
ASSOCIATE PROFESSOR
DEPARTMENT OF MICROBIOLOGY
BHARATHIDASAN UIVERSITY
TRICHY - 24

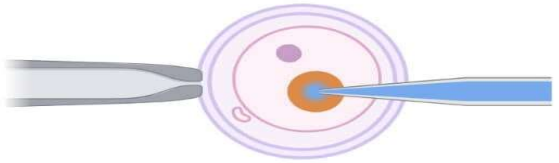




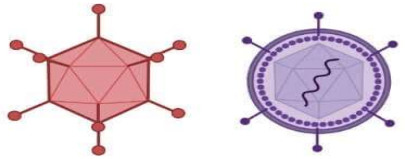


Gene Delivery Methods

Electroporator



Microinjection

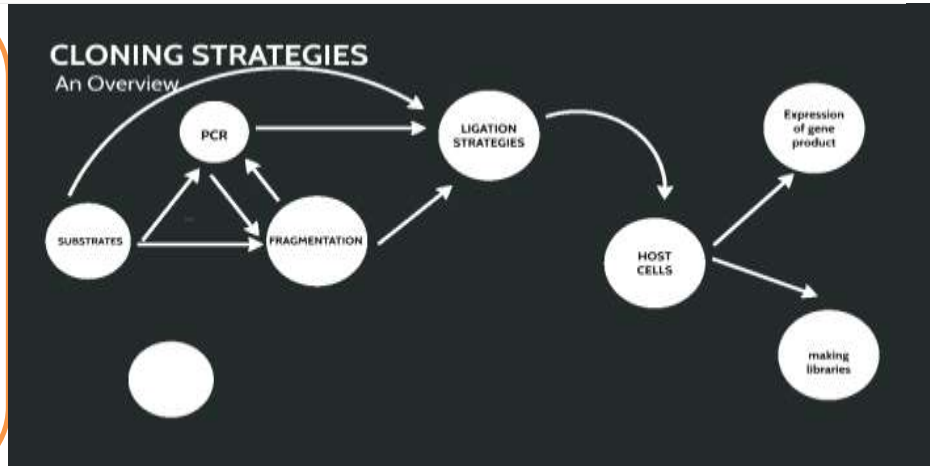


Biological (Viruses)

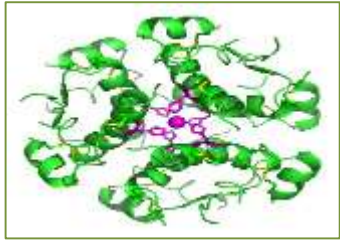
Chemical (Polymer and Liposome)

Gene Gun

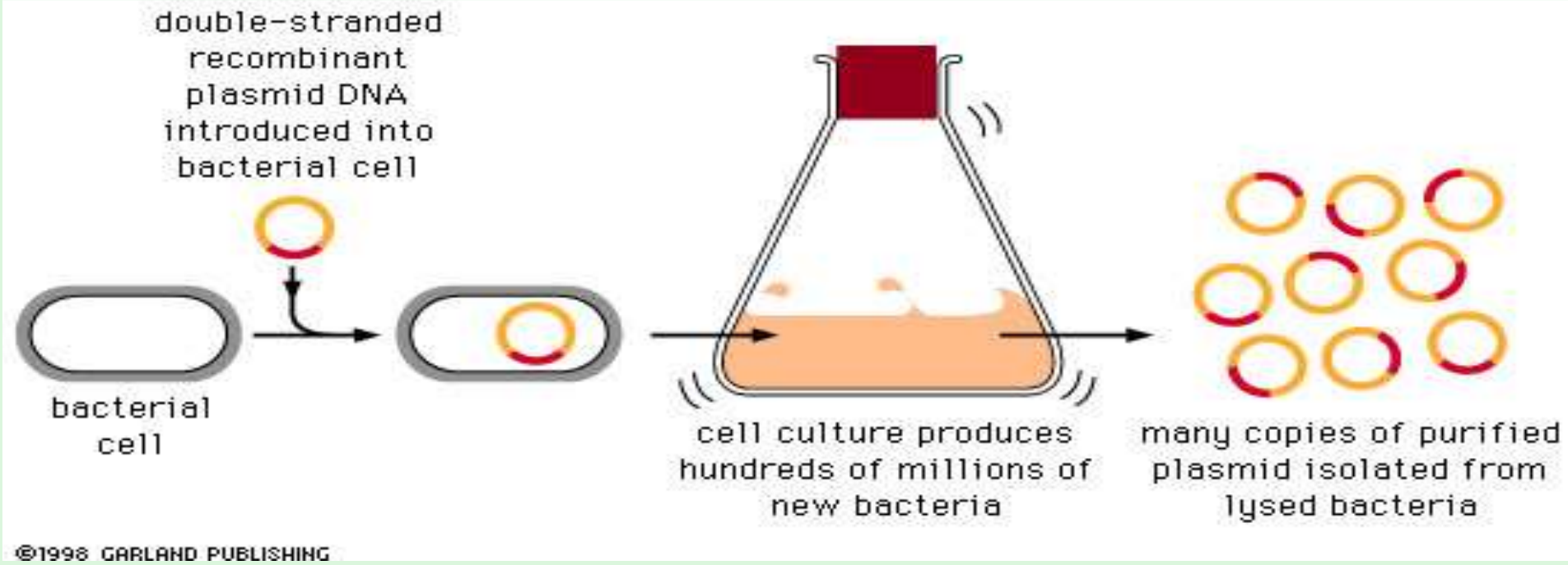
SURESH S S RAJA
DEPARTMENT OF MICROBIOLOGY
BHARATHIDASAN UIVERSITY
TRICHY - 24

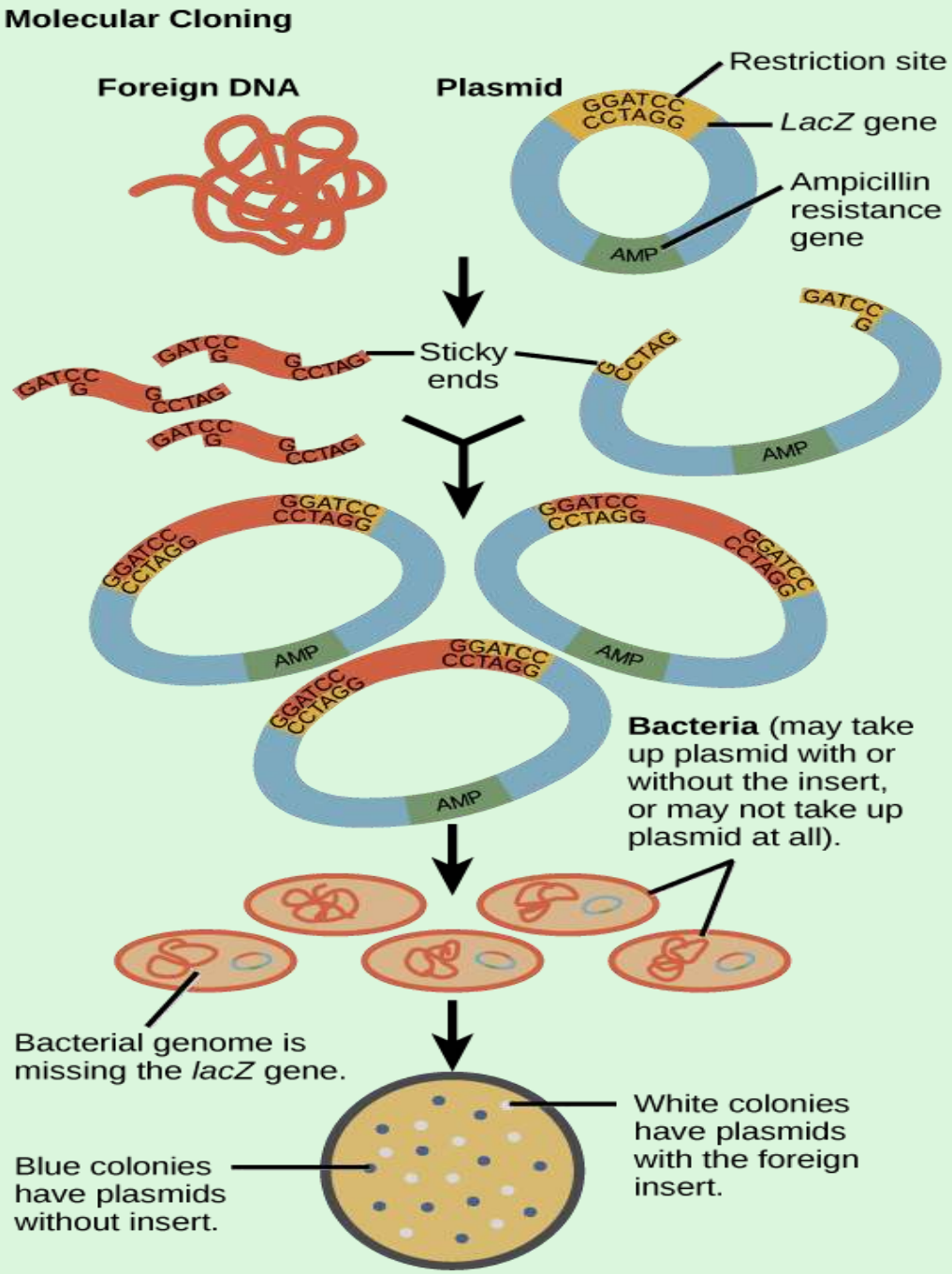


Molecular Cloning



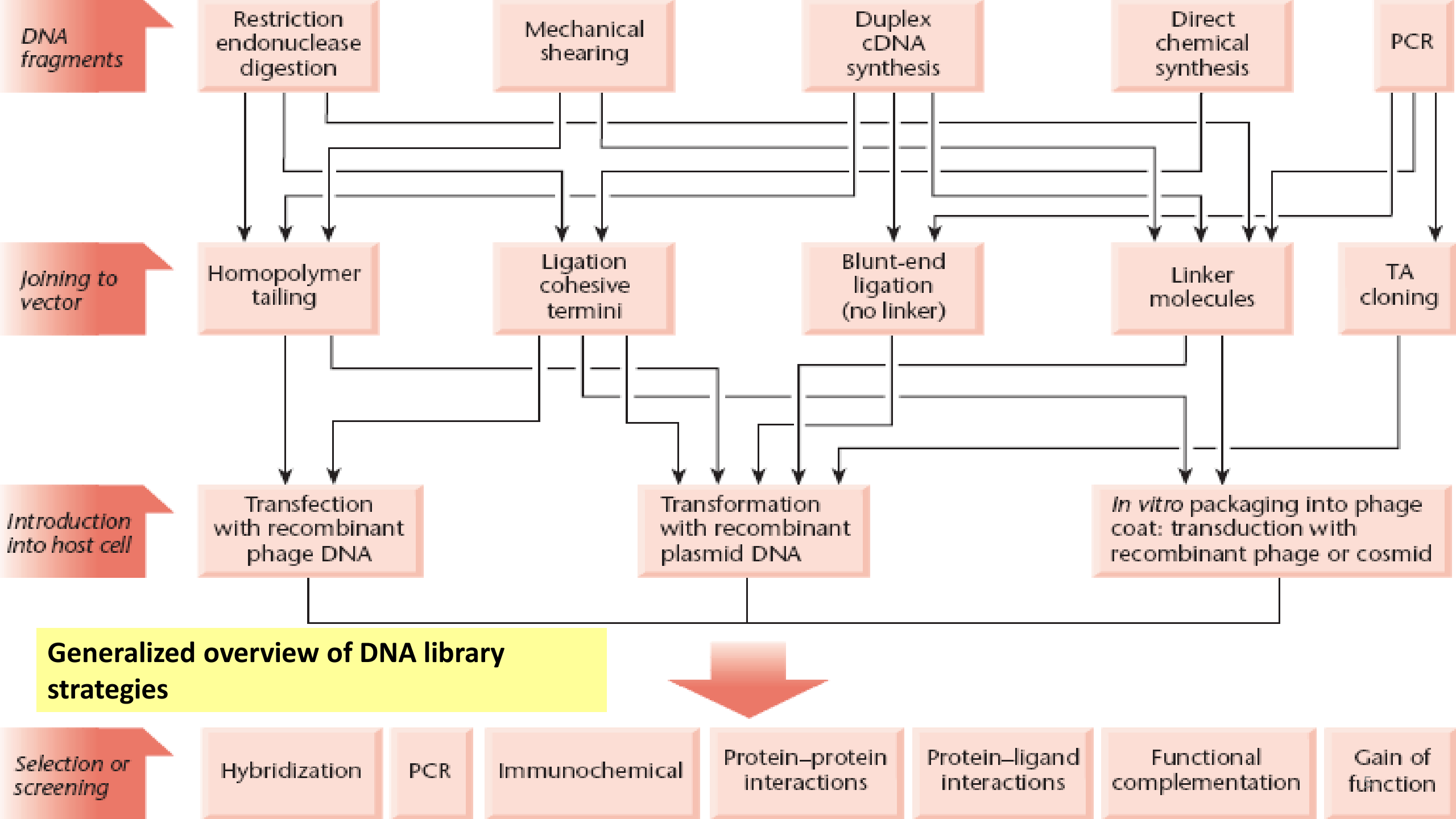
- Molecular cloning is a technique **to assemble recombinant DNA molecule and to amplify them in host organism**.
- Amplification: DNA fragment containing genes, DNA sequences (promoter, noncoding sequence, random DNA fragments)





- **6 steps in molecular cloning**

1. Choice of DNA vector and host organism
2. Preparation of cloning vector
3. Preparation of DNA to be cloned
4. Creation of recombinant DNA by ligation
5. Introduction of recombinant DNA into host organism
6. Screening for clones with desired DNA inserts and biological properties



Generalized overview of DNA library strategies

DNA libraries - collection of DNA sequences

- A DNA **library is a collection of clones of** DNA designed so that there is a high probability of finding any particular piece of the source DNA in the collection
- DNA libraries, like conventional libraries, are used to collect and store information
- All DNA libraries are collections of DNA fragments that represent a particular biological system of interest

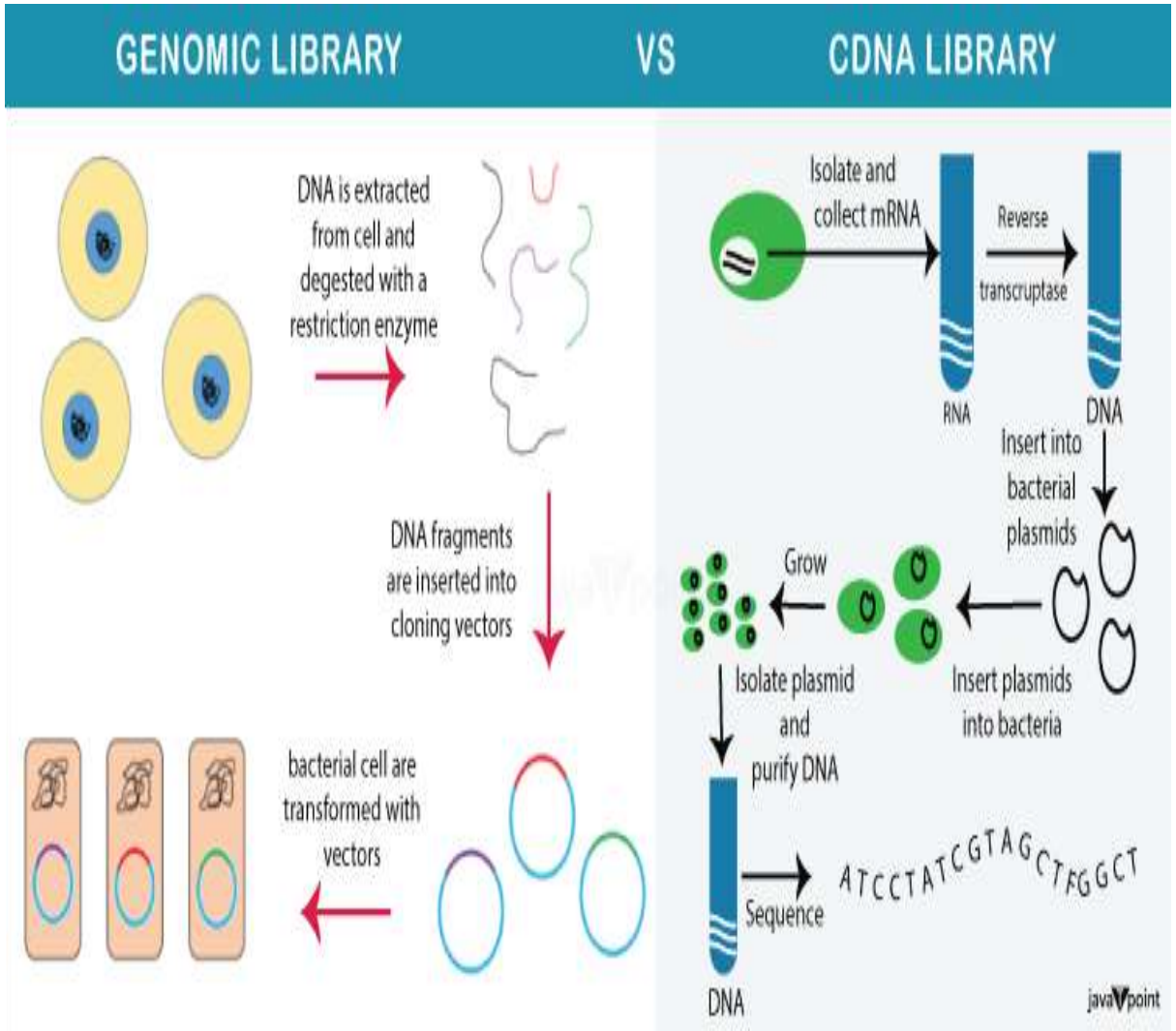
Uses of DNA libraries

To obtain the sequences of genes for analysis, amplification, cloning, and expression.

Once the sequence is known probes, primers, *etc*. can be synthesized for further diagnostic work using, for example, hybridization reactions, blots and PCR.

Knowledge of a gene sequence also offers the possibility of gene therapy.

Also, gene expression can be used to synthesize a product in particular host cells, *e.g*. synthesis of human gene products in prokaryotic cells.



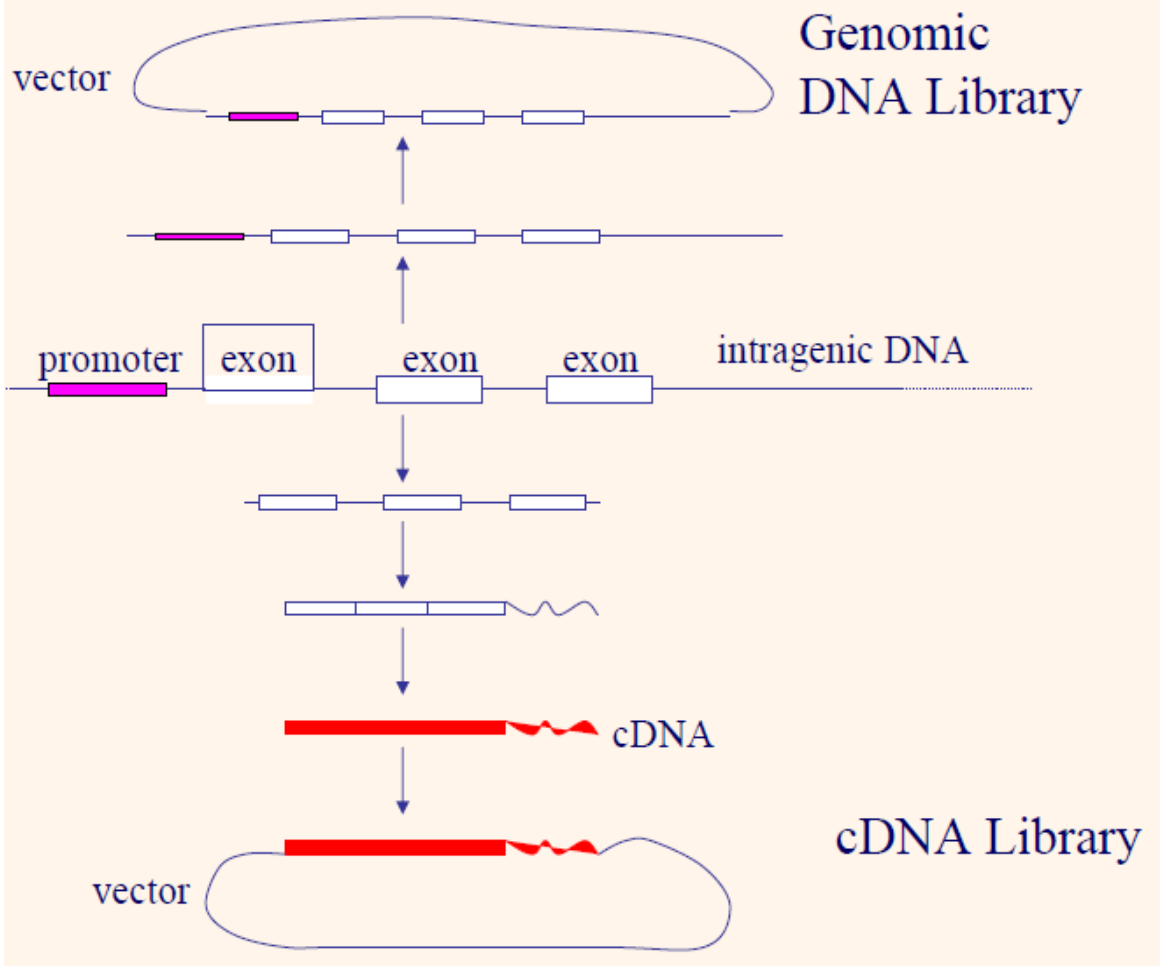
APPLICATIONS OF GENOMIC LIBRARY

1. Genomic library construction is the first step in any DNA sequencing projects.
2. Genomic library helps in identification of the novel pharmaceutically important genes.
3. Genomic library helps in identification of new genes which were silent in the host.
4. It helps us in understanding the complexity of genomes.
5. Serving as a source of genomic sequence for generations of transgenic animals through genetic engineering.
6. Study of the function of regulatory sequences in vitro.
7. Create cDNA libraries to determine what genes are being expressed at a particular time.

Types of DNA Libraries

- The **genomic library** contains DNA fragments representing the entire genome of an organism

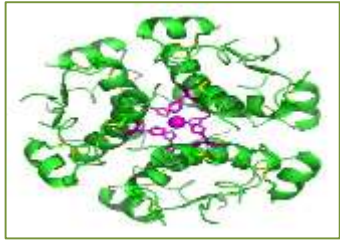
- The **cDNA library** contains only complementary DNA molecules synthesized from mRNA molecules in a cell

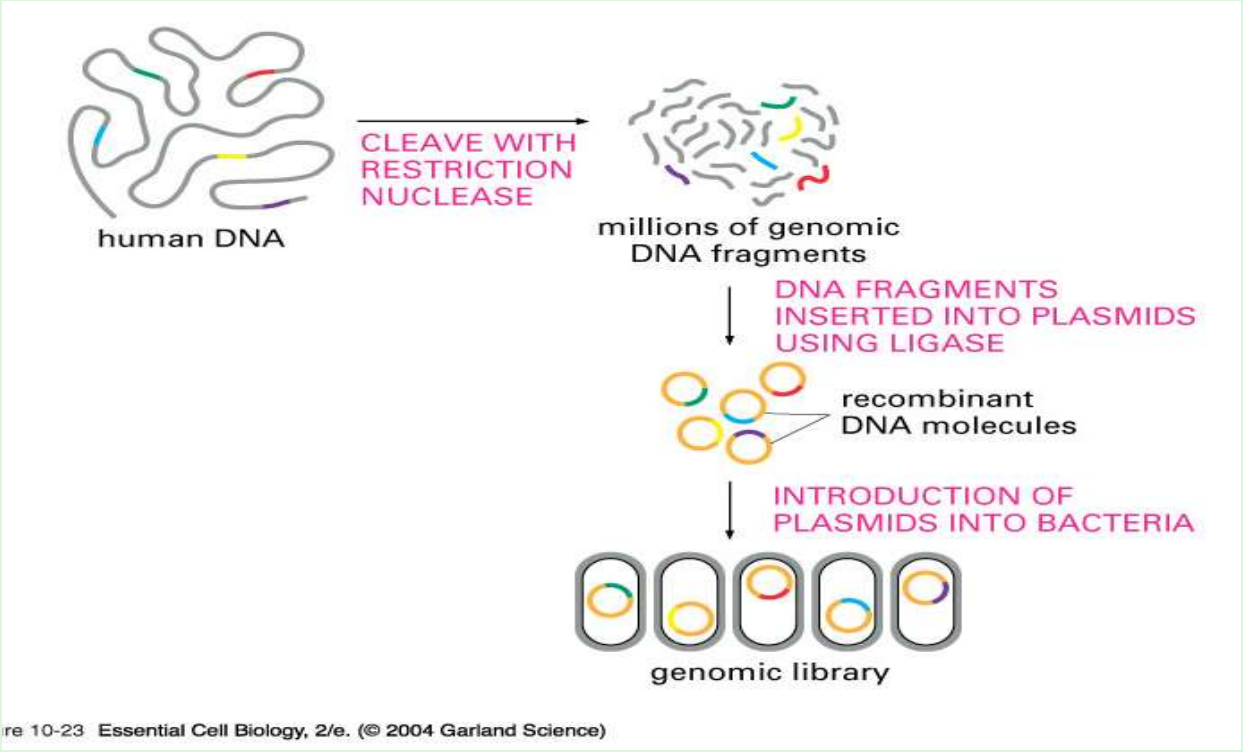


Genomic and cDNA library

Genomic Libraries	cDNA Libraries
• from genomic DNA	• reverse transcription of mRNA
• frequency of hits independent of gene expression levels	• dependent
• may contain promoters and introns	• no promoters or introns
• cannot express in heterologous system	• expression is feasible if linked to a suitable promoter
• useful for genome analysis, map-based cloning, promoter studies, etc	• useful for analysis of coding regions and gene functions

Genomic library





re 10-23 Essential Cell Biology, 2/e. (© 2004 Garland Science)

- Genomic library - a set of clones that represents the entire genome of a given organism.
- The number of clones that constitute a genomic library depends on
  - the size of the genome
  - the insert size for the particular cloning vector system.

For most practical purposes, the tissue source of the genomic DNA is unimportant because each cell of the body contains virtually identical DNA (with some exceptions).

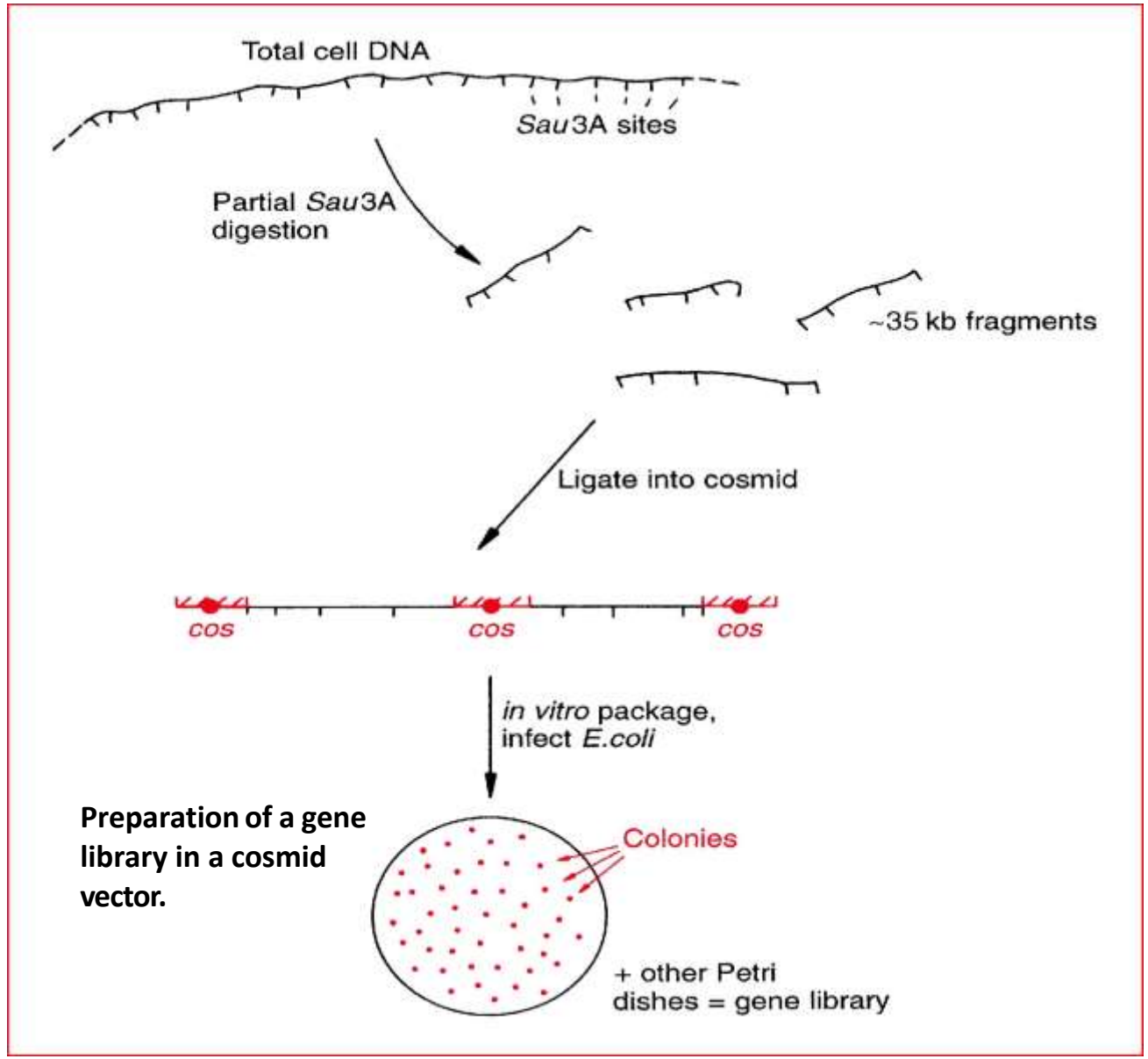
Genomic library

- **Genomic libraries can be retained for many years,and propagated so that copies can be sent from one research group to another.**
- It contains all DNA sequences such as :

1. **expressed genes**
2. **non-expressed genes**
3. **exons and introns**
4. **promoter and terminator regions and**
5. **intervening DNA sequences.**

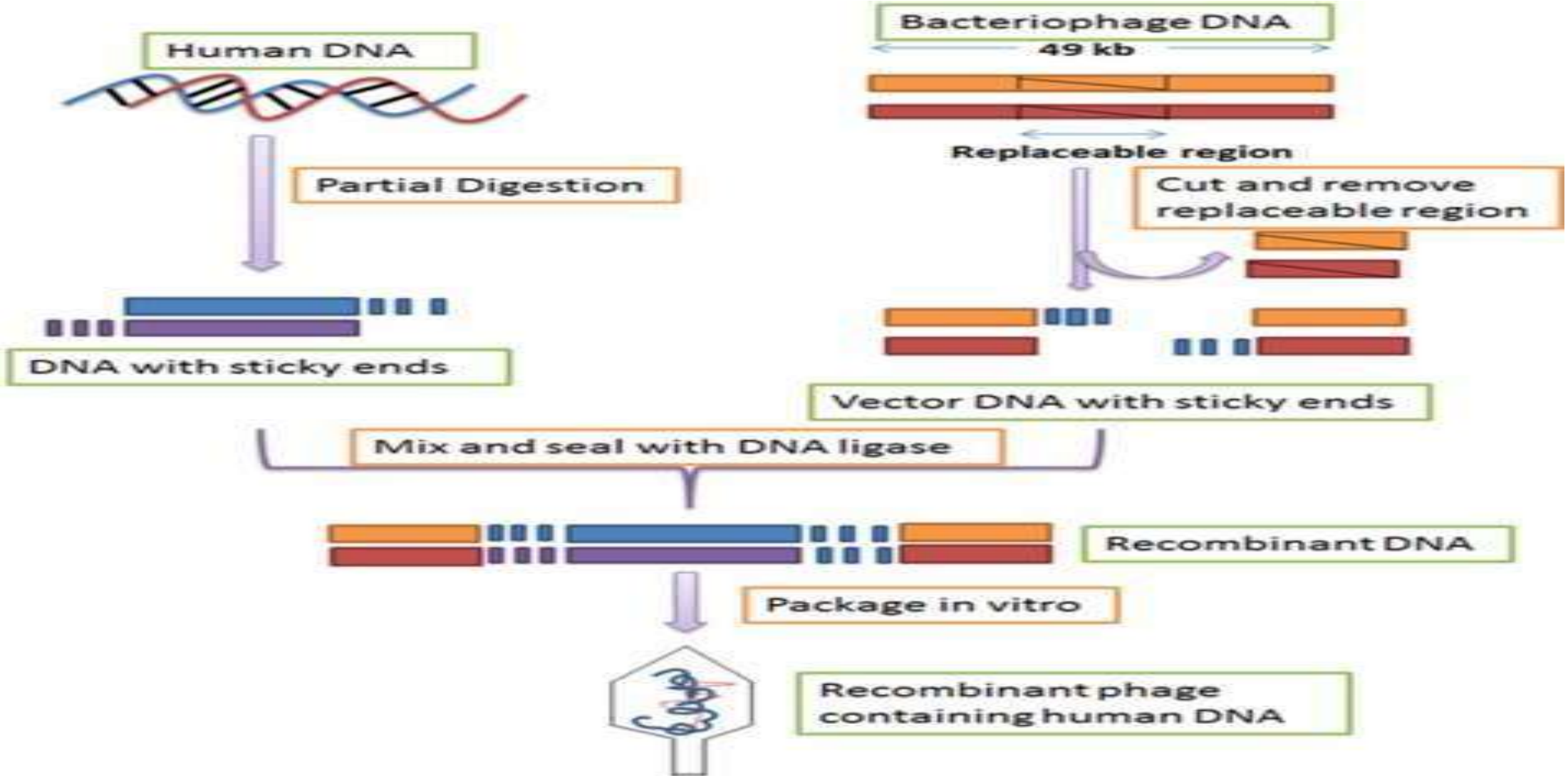
Construction of genomic library

- Construction of a genomic DNA library involves
  - **isolation,**
  - **purification and fragmentation of genomic DNA**
  - **followed by cloning of the fragmented DNA using suitable vectors.**
- **The eukaryotic cell nuclei are purified by digestion with protease and organic (phenol-chloroform) extraction.**
- The derived genomic DNA is too large to incorporate into a vector, and needs to be broken up into desirable fragment sizes.
- **Fragmentation of DNA can be achieved by physical method and enzymatic method.**
- The library **created contains representative copies of all DNA fragments present within the genome.**



Preparation of a gene library in a cosmid vector.

ConstructION Genomic library



Cloning of Genomic DNA

- Various vectors are available for cloning large DNA fragments.
- Vectors for larger DNA : λ phage, yeast artificial chromosome(YAC), bacterial artificial chromosome(BAC), etc.
- λ replacement vectors for construction of genomic library : λ DASH & EMBL 3
- **T4 DNA ligase** is used to ligate the selected DNA sequence into the vector.

Vectors for cloning of genomic DNA

1.High-capacity vectors

- The high capacity cloning vectors used for the construction of genomic libraries are cosmids, Bacterial artificial chromosomes (BACs), P1-derived artificial chromosomes (PACs) and Yeast artificial chromosomes (YACs).
- They are designed to handle longer DNA inserts, much larger than for λ replacement vectors.
- So they require lower number of recombinants to be screened for identification of a particular gene of interest.

2. λ replacement vectors

- The *λ EMBL* series of vectors are widely used for genomic library construction.
- The multiple cloning sites of these vectors flanking the stuffer fragment contain opposed promoters for the T3 and T7 RNA polymerases.
- The restriction digestion of the recombinant vector generates short fragments of insert DNA left attached to these promoters.
- This generates RNA probes for the ends of the DNA insert.
- These vectors can be made conveniently, directly from the vector, without recourse to sub-cloning.

Blunt end ligation

- Mainly three methods can be used to put the correct sticky ends onto the DNA fragments-

1. Cloning foreign DNA by adding linkers
2. Cloning foreign DNA by adding adaptors
3. Homopolymeric tail adding by using Terminal transferase enzyme.

Generating Ends by….

Linkers

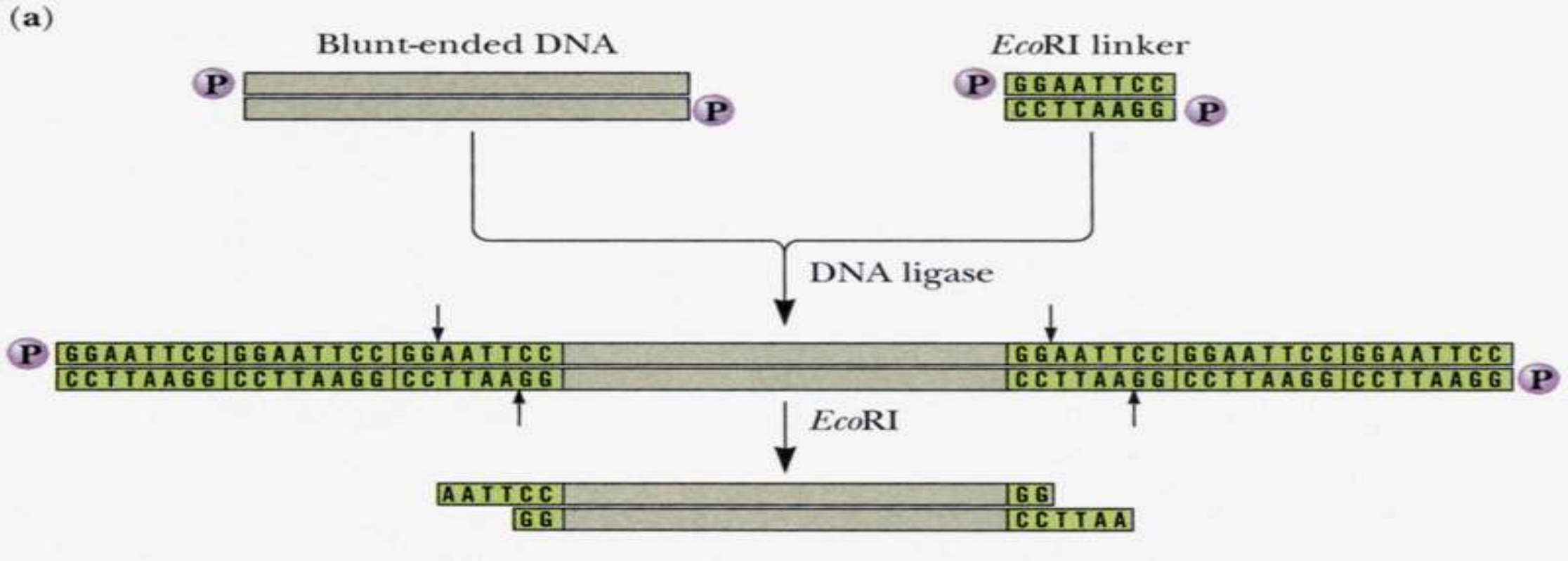
- Linkers are short stretches of double stranded DNA of length 8-14 bp that have recognition site for restriction enzymes.
- Linkers are ligated to blunt end DNA by ligase enzyme.
- The linker ligation is more efficient as compared to blunt-end ligation of larger molecules because of the presence of high concentration of these small molecules in the reaction.
- The ligated DNA can be digested with appropriate restriction enzyme generating cohesive ends required for cloning in a vector.
- The restriction sites for the enzyme used to generate cohesive ends may be present within the target DNA fragment which may limit their use for cloning.

Adapters

- These are short stretches of oligonucleotide with cohesive ends or a linker digested with restriction enzymes prior to ligation.
- Addition of adaptors to the ends of a DNA converts the blunt ends into cohesive ends.

1. **Linker**

- Linkers are the chemically synthesized double stranded DNA oligonucleotides containing on it one or more restriction sites for cleavage by restriction enzymes, e.g. Eco RI, Hind III, Bam HI, etc.
- Linkers are ligated to blunt end DNA by using DNA ligase.
- Both the vector and DNA are treated with restriction enzyme to develop sticky ends.
- The staggered cuts i.e. sticky ends are then ligated with T4 DNA ligase with very high efficiency to the termini of the vector and recombinant plasmid DNA molecules are produced.



Before cloning of a particular sequence in PCR, a primer, associated with a linker is used. This type of primer is called as linker-primer.

Now-a-days, two different linkers are used which has different RE sites with F/R primer. This strategy helps in directional cloning.

Limitations

- It may be the case that the restriction enzyme used to generate the cohesive ends in the linker will also cut the foreign DNA at internal sites.

Solution: CHOOSE ANOTHER RESTRICTION ENZYME

But there may not be a suitable choice if the foreign DNA is large and has sites for several restriction enzymes.

Methylation of internal restriction sites with the Appropriate modification methylase for example EcoRI methylase.

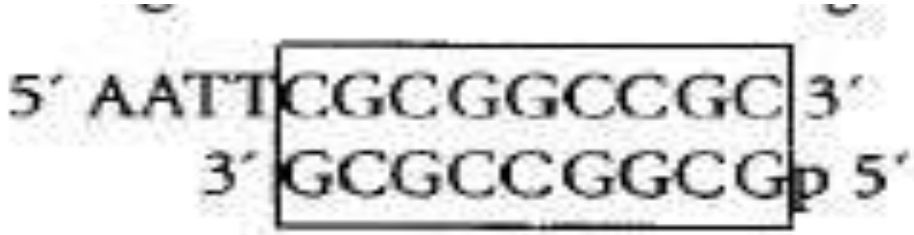
2. **Adaptors**

- They are also short double stranded oligonucleotides that carry an internal RE sites and single stranded tails at one or both ends.

- This protruding sequences can be ligated to DNA fragments containing a complementary single stranded terminus.

- After ligation, the DNA can be cleaved with appropriate RE to create new protruding terminus.

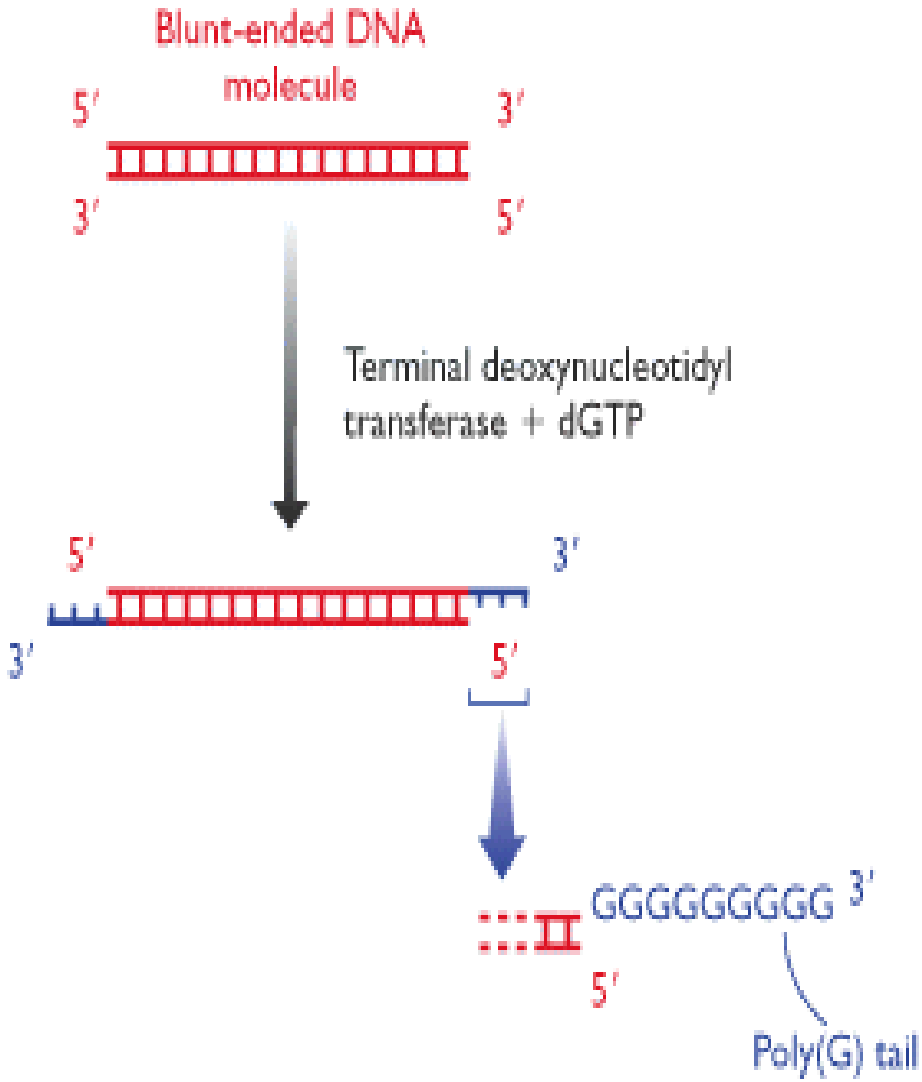
- Adaptors are available in two basic designs and a variety of specifications.
- 1. Some consists of a partial duplex formed between two oligonucleotides of different length; for example, the EcoR1-Not1 adaptor.



- 2. Another class of adaptor is supplied as an unphosphorylated single oligonucleotide whose sequence is partially self complementary. As an example- EcoR1-Pst1 adaptor.

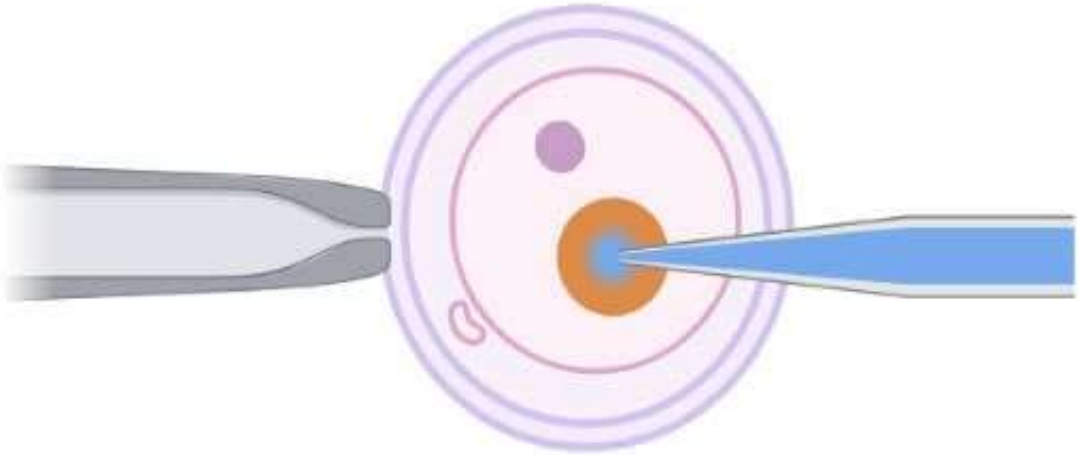
3. **Homopolymeric tailing**

- It is a technique by which sticky ends can be produced on a blunt-ended DNA molecule.
- In a homopolymer, all the subunits are same. A DNA strand made up entirely of deoxyguanosine is an example of homopolymer, and is referred to as polydeoxyguanosine or poly(dG).
- Tailing involves using the enzyme terminal deoxynucleotidyl transferase to add a series of nucleotides on to the 3'-OH termini of a double-stranded DNA molecule.
- The reaction when carried out in the presence of just one deoxynucleotide, then a homopolymer tail will be produced.

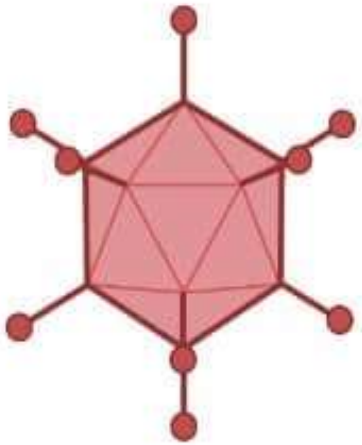


Gene Delivery Methods

Electroporator



Microinjection

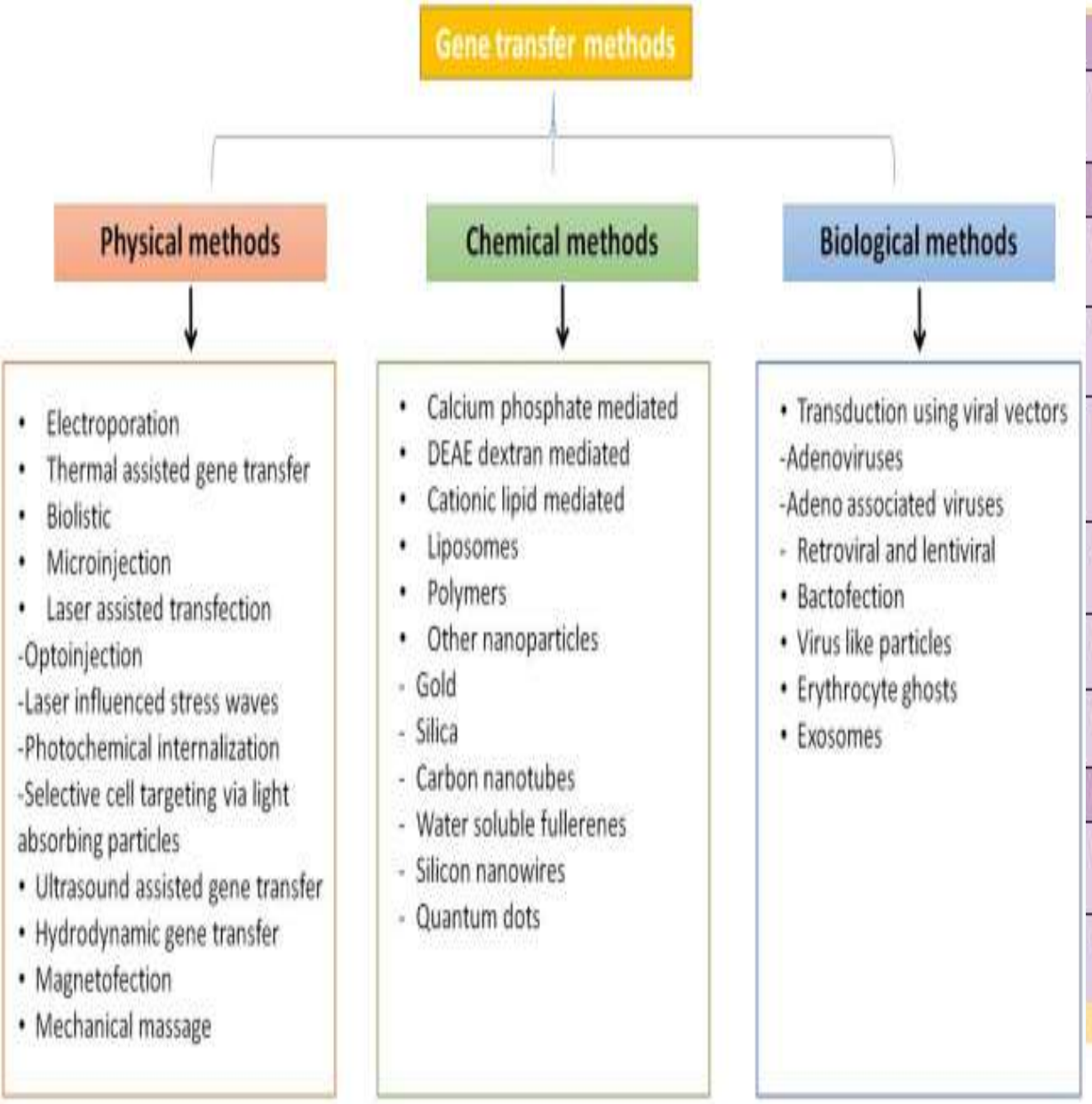


Biological (Viruses)

Chemical (Polymer and Liposome)

Gene Gun

GENE TRANSFER METHODS



	Advantages	Disadvantages
Naked DNA	No special skills needed Easy to produce	Low transduction efficiency Transient gene expression
Physical methods		
Microinjection	Up to 100% transduction efficiency (nuclear injection)	Requires highly specialized skills for delivery Limited to ex vivo delivery
Gene gun	Easy to perform Effective immunization with low amount of DNA	Poor tissue penetration
Electroporation	High transduction efficiency	Transient gene expression Toxicity, tissue damage Highly invasive
Sonoporation	Method well tolerated for other applications	Transient gene expression Toxicity not yet established
Laser irradiation	Can achieve 100% transduction efficiency	Special skills and expensive equipment necessary
Magentofection	Safety of method established in the clinic	Poor efficiency with naked DNA
Chemical methods		
Liposomes	Easy to produce Fusion liposomes improve transduction efficiency	Transient gene expression Toxicity, mildly immunogenic
Cationic polymers	Easy to manipulate for targeting	Transient gene expression Toxicity, mildly immunogenic

Source: From Mehier-Humbert and Guy, 2005; Conwell and Huang, 2005; Miyazaki et al., 2006.

Table 6 Summary of non-viral methods used for gene transfer.

Genomic Library

Advantages

- Identification of a clone encoding a particular gene of interest.
- It is useful for prokaryotic organisms having relatively small genomes.
- Genomic libraries from eukaryotic organisms are very important to study the genome sequence of a particular gene, including its regulatory sequences and its pattern of introns and exons.

Disadvantages

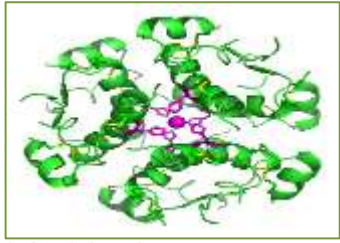
- Genome libraries from eukaryotes having very large genomes contain a lot of DNA which does not code for proteins and also contain noncoding DNA such as repetitive DNA and regulatory regions which makes them less than ideal.
- Genomic library from a eukaryotic organism will not work if the screening method requires the expression of a gene.

Applications

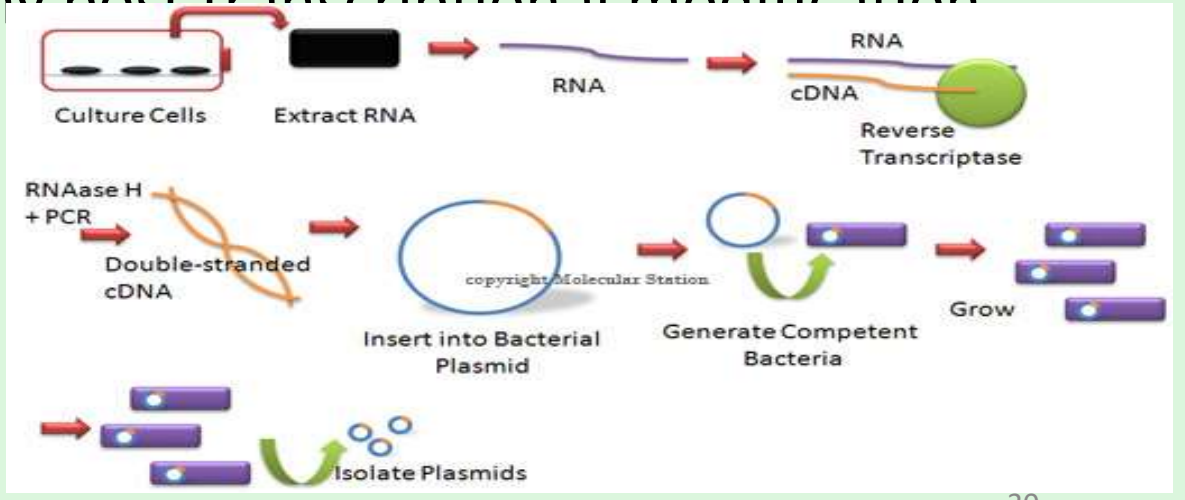
- To determine the complete genome sequence of a given organism.
- To study and generate transgenic animals through genetic engineering, serving as a source of genomic sequence.
- To study the function of regulatory sequences

  in vitro.
- To study the genetic mutations.
- Used for genome mapping, sequencing and the assembly of clone contigs.

cDNA library



- cDNA (complementary DNA) library - a collection of the mRNA purified from a particular source under specific condition (either a collection of cells, a particular tissue, or an entire organism), which has been converted back to a DNA template by reverse transcriptase.
- It thus represents the genes that were **being actively transcribed in that particular source under the physiological, developmental, or environmental conditions** when the mRNA was purified.
- cDNA libraries represent sequences after any post-transcriptional modification (e.g. Splicing of introns)
- The advantage of cDNA library is that it contains only the coding region of a genome



Advantages of cDNA

- The conversion of mRNA sequences to ds cDNA offers several advantages, which include:

  (i) mRNA is naturally labile single stranded molecule, and its conversion into more stable ds cDNA facilitates long-term storage of these sequences

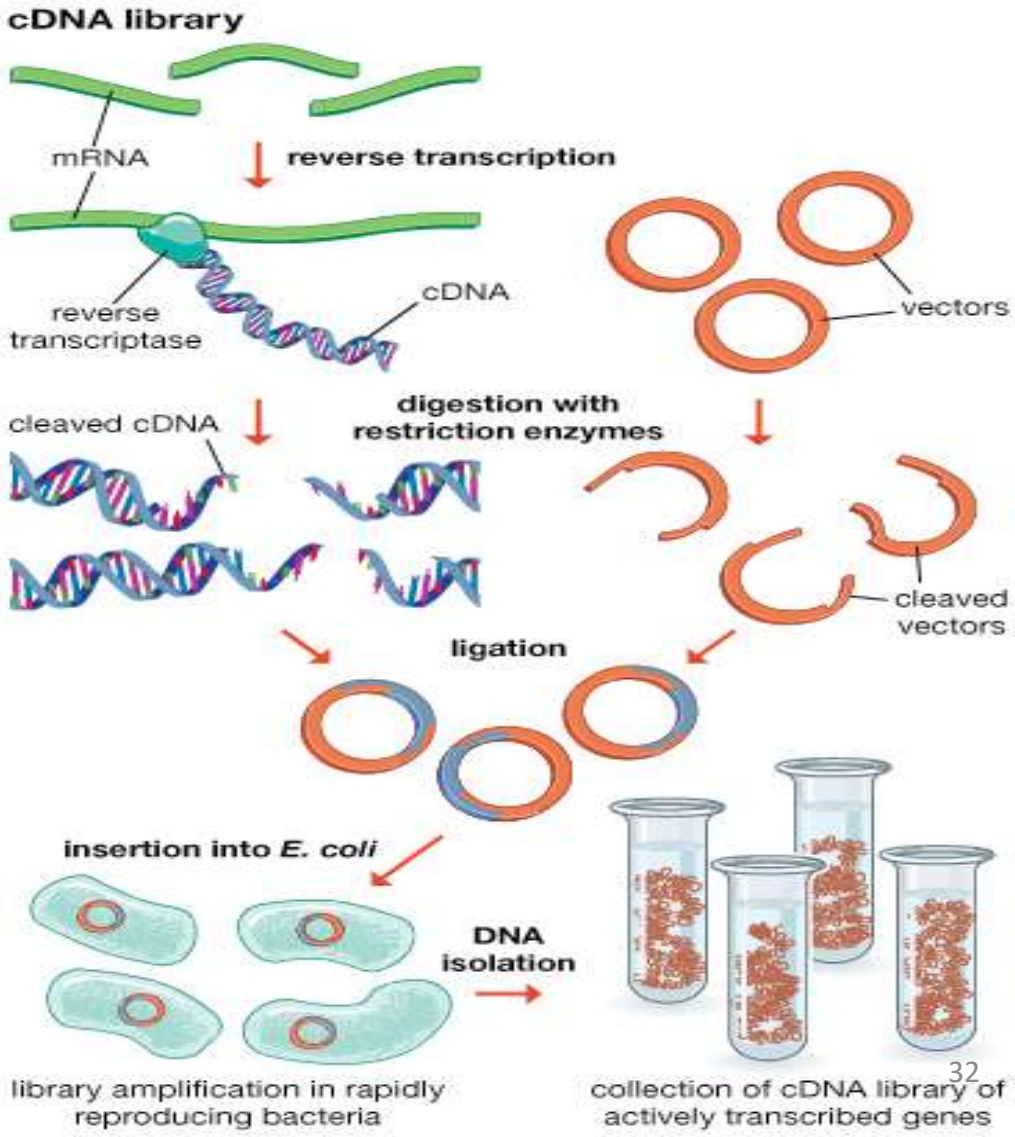
  (ii) cDNA is representative of the mRNA population (i.e., expressed genes) of the cell. Thus, the synthesis of cDNA is the creation of a permanent biochemical record of the cell at the time of lysis.

  (iii) By synthesizing and cloning the resulting cDNA from a single source, a method for propagating the cDNA is created. This approach is greatly facilitated by the huge variety of vectors compatible with an equally impressive variety of hosts

  (iv) cDNA molecules, both long and short, can be used to screen (hybridize to) members of much more complex

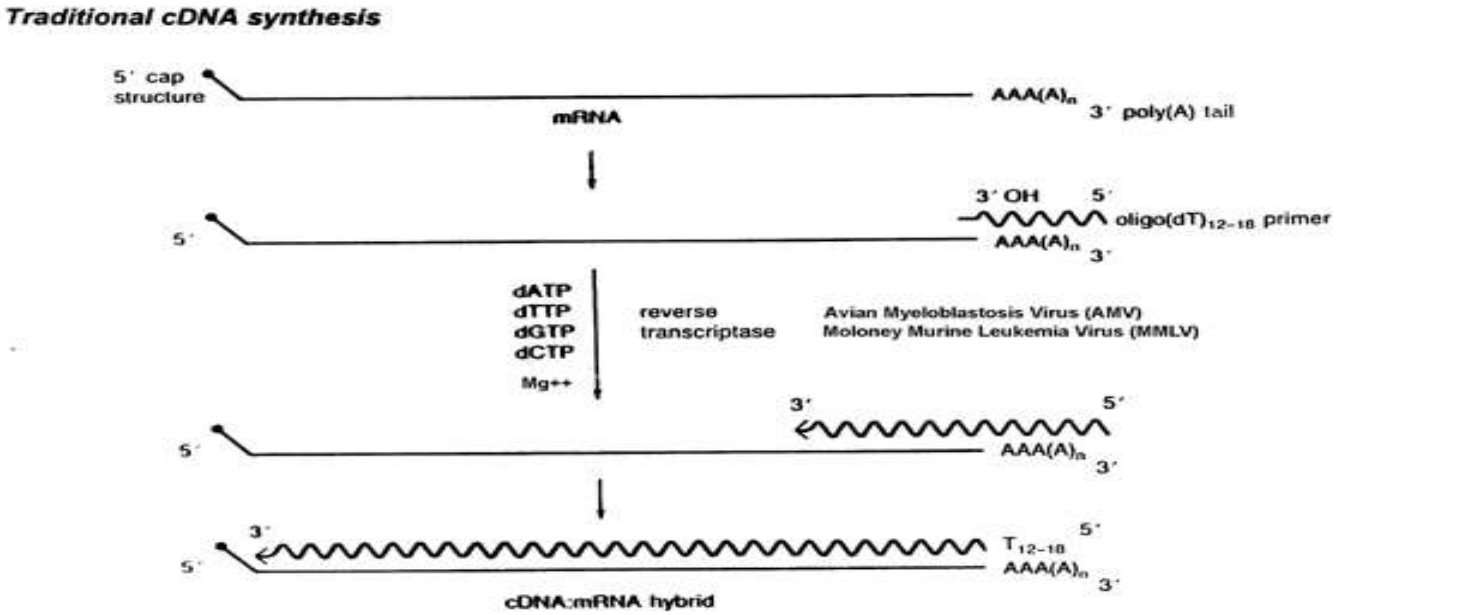
Construction of cDNA library

- Synthesis of first and second cDNA strands
- Methylation of cDNA at internal sites
- Attachment of linkers and adaptors, and restriction digestion
- Ligation of ds DNA to vectors
- Introduction into host systems for amplification of cDNA clones



First cDNA strand synthesis

- First strand cDNA synthesis using mRNA template is a reaction catalyzed by reverse transcriptase
- mRNA may be subjected to size fractionation before first cDNA strand synthesis
- Reverse transcriptase requires primer to extend on, and the reaction leads to the formation of a hybrid between single stranded cDNA and mRNA (i.e., ss cDNA:mRNA hybrid)



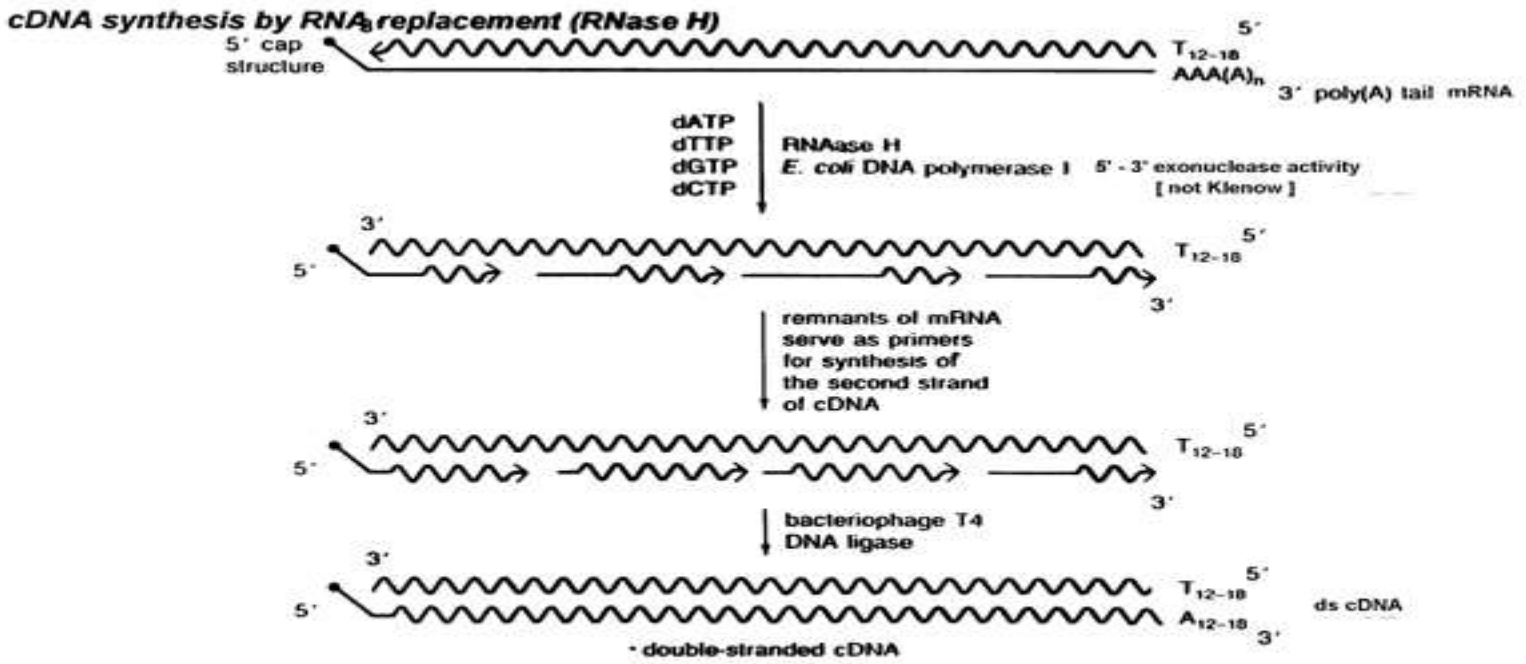
Synthesis of the first strand of cDNA using an oligo(dT) primer and reverse transcriptase.

Degradation of mRNA

- The mRNA template in ss cDNA:mRNA hybrid is degraded by alkali treatment or by RNase H-catalyzed nicking
- This step provides primers for second cDNA strand synthesis (not in case of self-priming of second cDNA strand and random primers catalyzed second strand synthesis)

Second cDNA strand synthesis

- Second cDNA strand synthesis using the first strand as template is a reaction catalyzed by a DNA-dependent DNA polymerase such as *E. coli* DNA polymerase I or Klenow enzyme.
- The reaction requires primers and the end result of this step is the formation of ds cDNA



Replacement synthesis of double-stranded cDNA.

Effect of mRNA length on cDNA synthesis

- The length of specific mRNA, and hence the cDNA that is to be cloned, influences the choice of primer and vector
- The majority of mRNA species are between 1.5 and 2.0 kb in length. For these 'standard' mRNAs, the primer of choice is oligo (dT), which is 12-18 nucleotides in length
- The oligo (dT) primer will anneal to the poly (A) tract, which is found at the 3'-terminus of most eukaryotic mRNAs (histone is a notable exception)
- There are, however, a number of mRNAs, which are extremely long, and it would be difficult to synthesize complete cDNA copies from these long mRNAs in one piece

Primers used for first cDNA strand synthesis

- The first step in the synthesis of ds cDNA is the formation of a ss cDNA:mRNA hybrid, which results from the synthesis of first cDNA strand using mRNA as a template
- First cDNA strand synthesis, either total RNA or poly (A) RNA can be used depending on the purpose. If the cDNA is to serve as a template for PCR, total RNA is usually sufficient, while poly (A) RNA is employed if the cDNA is to be used as a hybridization probe
- For first cDNA strand synthesis, mRNA template and primers are added together
- The preparation is heated to 70°C to melt the secondary structure of a RNA and then cooled slowly to room temperature to allow primer hybridization (called annealing of primers)

Oligo (dT) Primers

- Oligo (dT) primers that represent a chemically synthesized stretch of 16-20 thymidinne residues are used when the mRNA species for which cDNA is required is polyadenylated at its 3'- end
- If the mRNA is not polyadenylated, oligo (dT) primers cannot be used
- The primer is added to the reaction mixture in large molar excess so that each molecule of mRNA binds several molecules of oligo (dT) primers at its long poly (A) tail
- The annealed primer supplies free 3'-OH group, which is extended by reverse transcriptase

Random primers

- Random primers (or random oligonucleotides) are a collection of chemically synthesized oligonucleotides of random sequence, usually hexamers
- These are produced by oligomerization of equal quantities of mixed A, G, C, and T residues, so that all possible hexameric sequences are present
- The hexamers bind throughout the length of the mRNA molecule (wherever complementary) resulting in internal priming of first cDNA strand instead of priming from the 3'-end of the template

Random primers

- Random primers, however, create certain problems, for example,

  (i) If one is interested in cDNA synthesis from poly (A) RNA, the product will be contaminated from cDNAs synthesized from poly (A) RNAs as well

  (ii) cDNA clones derived from random priming of first strand cDNA tend to be smaller than clones derived from oligo (dT) priming

  (iii) The efficiency of priming varies from one mRNA to another, depending on the size of mRNA, and complexity and composition of the random primers pool

  (iv) The primers have to be removed between first and second strand synthesis steps, otherwise second strand synthesis will also initiate by random priming

Oligo (dT) Primer-adaptor or Oligo (dT) Primer-linker

- For first cDNA strand synthesis oligonucleotides such as oligo (dT) primer-adaptor may also be employed
- These primers consist of oligo (dT) appended at their 5'-end with overhanging sequences that will support the addition of restriction enzyme sites or useful sequences on one end
- The purpose of such primer-adaptor is directional cloning of ds cDNA, hence such first cDNA strand synthesis is done in combination with second strand synthesis with a similar primer-adaptor having different restriction enzyme site
- Oligo (dT) primer-linker works on the same principle for directional cloning

Asymmetrically tailed plasmid (Oligomerically tailed plasmid)

- Synthesis of the first strand is also primed from an oligo (dT) tail covalently attached to one end of the linearized plasmid
- This strategy allows direct and efficient cloning of cDNA into the vector
- Being large in size, it is often not possible to incorporate enough oligomerically tailed plasmid in the reaction mixture to achieve a high primer:template ratio

Primers used for second cDNA strand synthesis

- In order to obtain a molecule that is stable, and can be manipulated and cloned into a vector, the mRNA in the ss cDNA:mRNA hybrid needs to be replaced with a cDNA strand of the same sequence
- This second cDNA strand synthesis is conveniently carried out by *E. coli* DNA polymerase I (Kornberg enzyme), Klenow enzyme, reverse transcriptase, or T4 DNA polyrnerase
- Thermostable DNA polymerases such as *Tth* and *Taq* DNA polymerases can also be used, albeit rarely

Self-priming

- This strategy takes advantage of the tendency of reverse tanscriptase to form a hairpin loop during the first cDNA strand synthesis, which is used as a primer for second cDNA strand synthesis
- This hairpin is formed due to spontaneous intramolecular base pairing during reverse transcriptase reaction. This technique is advantageous because a single primer that is used for first cDNA strand synthesis is used to synthesize the second cDNA strand as well
- This method leads to higher yields of ds cDNA
- For cloning ds cDNA into a suitable vector, hairpin loop is removed in a reaction catalyzed by S1 nuclease. As it is an uncontrolled reaction, there may be frequent loss of cDNA sequences corresponding to the 5'-end of mRNA

RNA fragments generated by RNase H treatment (replacement synthesis)

- RNase H recognizes a ss cDNA:mRNA hybrid, and makes many nicks or gaps in the RNA strand. Many RNA fragments thus created along the length of the first cDNA strand provide free 3'-OH groups to be used as primers for extension by polymerase. The reaction has three main advantages:

  (i) The priming is very efficient

  (ii) Such priming does not need any further treatment or purification, and can be carried out directly using the products of the first cDNA strand synthesis reaction and

  (iii) It eliminates the need to use S1 nuclease, and hence is free from disadvantages such as loss of cDNA sequences that results from S1 nuclease treatment

Oligonucleotide priming

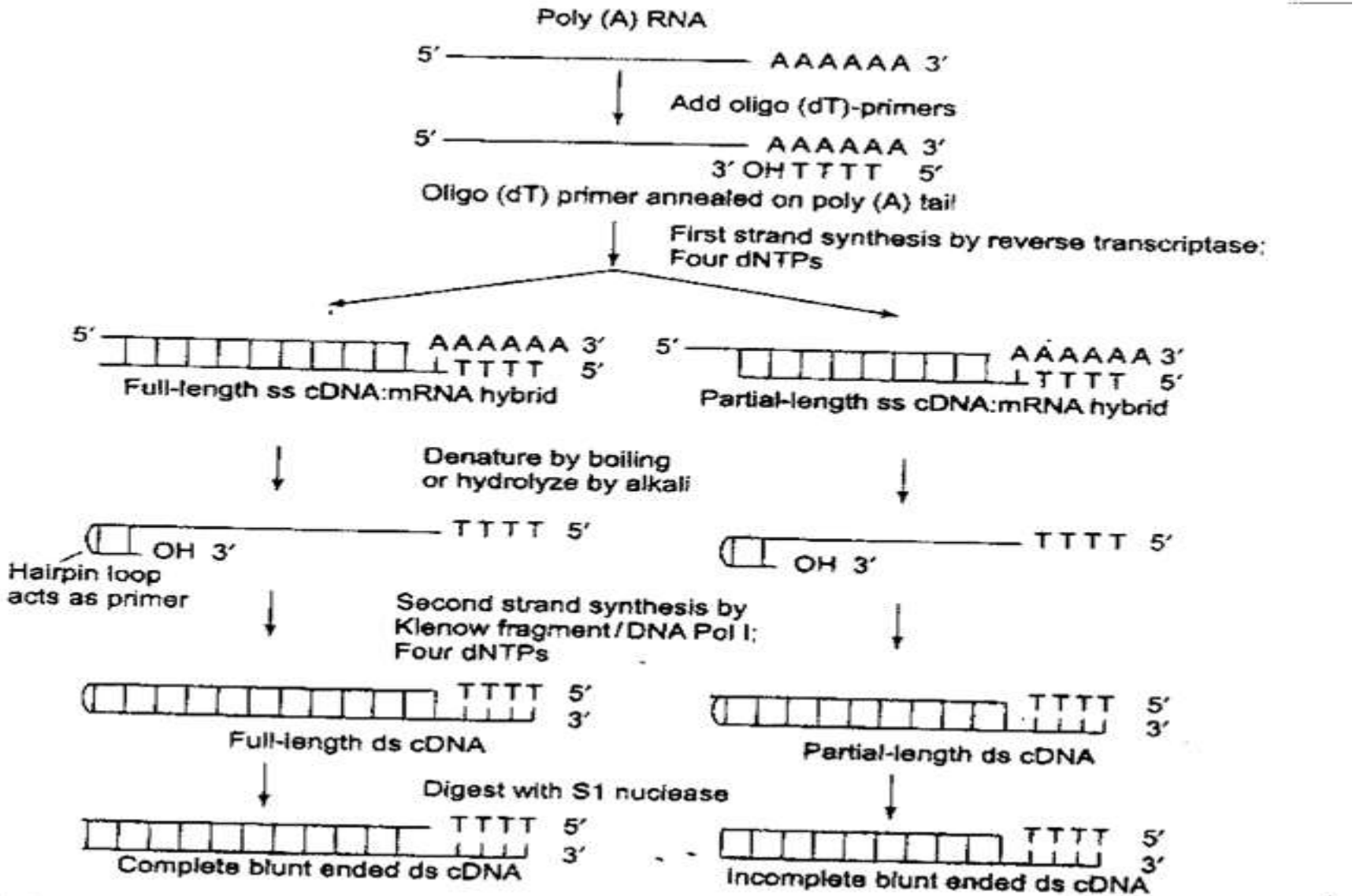
- In this method, homopolymeric tail is added to the 3'-end of the newly synthesized first cDNA strand
- This tailing is done by exploiting the activity of terminal deoxynucleotidyl transferase (TdT), a template independent DNA polymerase from calf thymus, in the presence of only single type of dNTP (take for example, addition of oligo dC tail by TdT in presence of dCTPs)
- This oligo dC tail then acts as an annealing site for a chemically synthesized complementary oligonucleotide primer (oligo dG)

Oligonucleotide-adaptor priming or Oligonucteotide-linker priming

- In this method, once first strand is synthesized using oligo (dT) primer-adaptor, a homopolymeric tail is added to its 3'-end in a TdT-catalyzed reaction in the presence of dCTP
- First cDNA strand synthesis, oligo (dT) anneals to poly (A) tail, and the adaptor provides a recognition site for a restriction endonuclease
- This oligo (dC) tail provides complementary sequences for annealing of oligonucleotide primer
- The mRNA strand in a ss cDNA:mRNA hybrid formed after first cDNA strand synthesis is subjected to alkali hydrolysis or denaturation

Random priming

- Similar to first cDNA strand synthesis by random primers, second cDNA strand synthesis by Klenow enzyme can also be primed by random primers
- Random priming of the second cDNA strand is done after alkali digestion of RNA in ss cDNA:mRNA hybrid
- Random priming prevents 3' or 5' bias in the resulting library, but with this method full-length cDNA is hardly obtained



Self priming classical method of ds DNA synthesis

Enrichment methods

- The library is expected to contain cDNAs for abundant mRNAs, such as general ‘house keeping’ genes and for those genes whose expression is specific to that particular tissue
- If cDNA encoding a specific protein is needed, it is wise to use a cDNA library from a tissue with a high quantity of RNA for that protein relative to other RNAs
- For example in the chicken oviduct, mRNA-encoding ovalbumin is superabundant, hence the starting population is naturally so enriched in ovalbumin mRNA

Construction of eukaryotic cDNA libraries in expression vectors

- **Eukaryotic cDNA that encodes functional protein is cloned in particular type of expression vector**
- **Most eukaryotic expression vectors are shuttle vectors that contain:**

  - **Promoter**
  - **Two origins of replication**
  - **eukaryotic transcriptional and translational stop signals**
  - **A sequence that enables polyadenylation of mRNA**
  - **Two selectable marker gens, one for *E. coli* and other for eukaryotic host cells**

cDNA libraries generated by RT-PCR

- RT-PCR can be used to provide the cDNA for library construction when the source of mRNA is unsuitable (e.g. very small amount) for conventional approaches

- In this technique instead of gene-specific primers, universal primers can be used

C-DNA Library

Advantages

- It is enriched with fragments from actively transcribing genes.
- Introns do not interrupt the cloned sequences. Introns would pose a problem when their goal is to produce eukaryotic protein in bacteria, since they lack mechanism of removing introns.
- cDNA maybe inserted into vectors and then cloned.

Disadvantages

- cDNA libraries contain only the parts of genes found in mature mRNA. However, the sequences before and after the gene,

  for example, those involved in the regulation of gene expression, will not occur in a cDNA library.
- Construction of a cDNA library cannot be used for isolating the genes expressed at low levels as there will be very little mRNA for it in any cell type and may completely be out manoeuvred by the more abundant species

Applications of C-DNA library

- Discovery of novel genes.
- *in vitro* study of gene function by cloning full-length cDNA.
- Determination of alternative splicing in various cell types/tissues.
- They are commonly used for the removal of various non-coding regions from the library.
- Expression of eukaryotic genes in prokaryotes as they lack introns in their DNA and therefore do not have any enzymes to cut it out in transcription process. Gene expression required either for the detection of the clone or the polypeptide product may be the primary objective of cloning.
- To study the expression of mRNA.

Comparison

Genomic library	Feature	C-DNA library
Ideally, all genomic sequences	**Sequences present**	Structural genes are trancribed
No No	**Contents affected by :** **(a) Developmental stage** **(b) Cell type**	Yes Yes
	Features of DNA insert(s) representing a gene:	
As present in genome	**(a) Size**	Ordinarlly smaller
Present	**(b) Introns**	Absent
Present	**(c) 5′- and 3′- regulatory sequences**	Absent
	As compared to the genome	
In amplified genomic libraries	**(a) Enrichment of sequences**	For abundant mRNAs
In amplified libraries	**(b) Redundancy in frequency**	For rare mRNA species
Not possible	**(c) Variant forms of a gene**	For such genes,whose RNA transcripts are alternatively spliced