Genomic DNA & cDNA Libraries

Debdeep Dasgupta, PhD

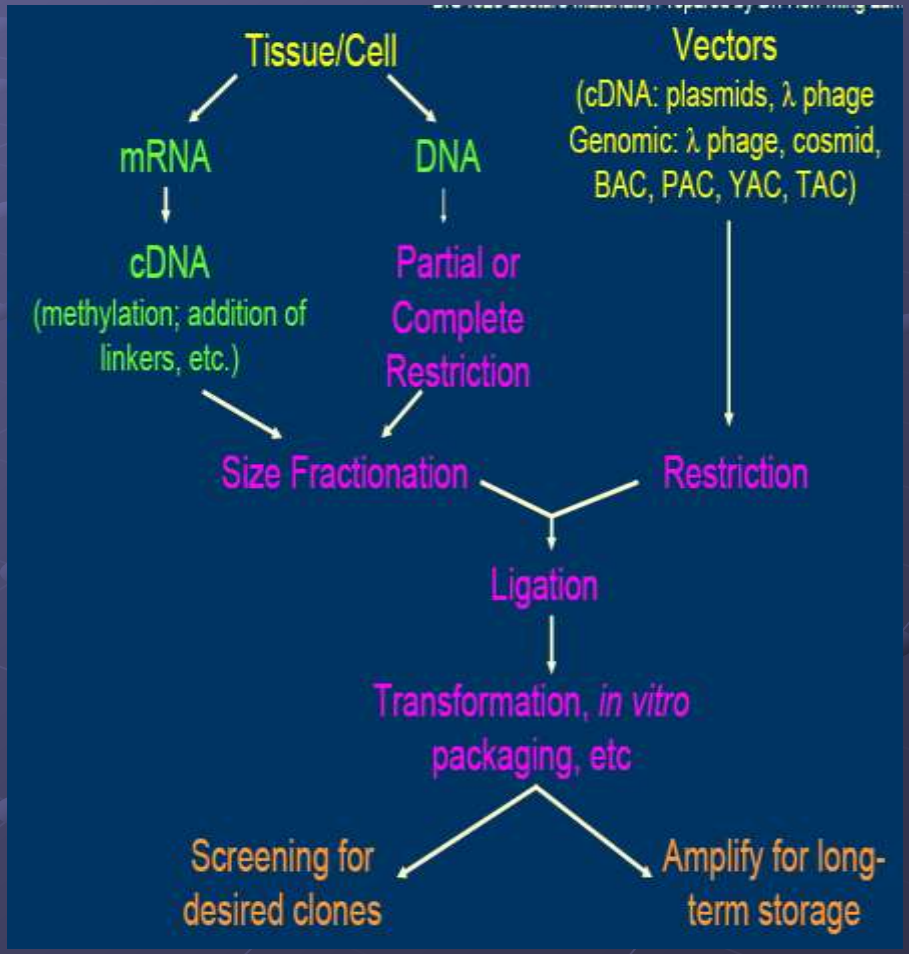
Dept. of Microbiology, Surendranath College

Introduction

- The use of genetic information is a powerful tool that today is becoming more readily available to scientists.
- In order to use this powerful tool it necessary to know how to navigate throughout the entire genome. The human genome is about 3 x 10E9 bp.
- In humans this project is known as Human Genome Project.

How to generate Genomic information

- There two ways in which genomic information is obtained.
- Genomic library which contains the entire human Genome (exons and introns)
- cDNA (complementary DNA) library the contains only expressed genomic information (only exons)



- Lambda was first discovered at the Pasteur Institute by Andre Lwoff when he observed strains of E. Coli.
- He showed that the cells of these bacterial strains carried bacteriophage in a dormant form (prophage).
- Phage can alternate between lysogenic (non-productive) and lytic (productive) growth cycles.

λ Bacteriophage

- Double stranded DNA molecule
- 5' twelve-base-pair sticky ends (cos sites)
- It is used as a cloning vector, accommodating fragments of DNA up to 15 kilobase pairs long. For larger pieces, the cosmid or YAC's are used.
- Will accept foreign DNA and still complete their life cycle.
- Distinguish cells that have foreign and non foreign DNA.
- Should replicate in host
- Gene of interest can be identified and grown in large amounts.
- Non essential genes can be removed and replaced by foreign gene.

Cont.

- Should carry one or more selectable markers that identify the parent and recombinant vectors
- Should have restriction sites in non-essential regions of DNA into which foreign DNA can be inserted
- easy to make and maintain library

Enzymes needed:

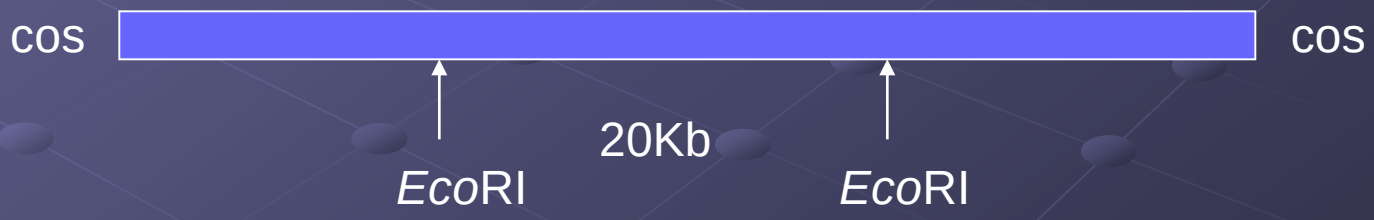
1. Restriction enzymes: cuts the DNA at specific sequences to generate a set of fragments

2. DNA ligase: inserts DNA restriction fragments into replicating DNA molecules to produce recombinant DNA

Lambda vectors

- 1) Insertion vectors

- 2) Replacement vectors



- http://www.sh.lsuhsc.edu/gradcore/IDSP117/16

limitations

- size of DNA to be introduced into the host cell

- Problem: when making genomic libarary of large size (plants and mammals) only a portion of those fragments will be represented. If gene of interest is located in a large fragment, then you won't be able to isolate that gene from the library.
- Solution :use a vector that can accept large fragments of DNA

Vector types:

1. **Plasmids**- small circular DNA molecules which can replicate their DNA independently of their bacterial chromosome. They are found naturally in bacteria and replicate inside the bacterial cell. They can insert pieces up to 10kb(kilobases) or 100 to 10,000 base pairs. Examples: pBR322 and pUC18
2. **Bacteriophage λ**– They are double stranded linear DNA vector. They replicate in E. Coli in the lytic or lysogenic mode. They can insert fragments up to 15kb. Examples are λgt10 and λZAP
3. **Cosmids-** are hybrid vectors of λ phage and plasmids. They can replicate their DNA in the cell with a plasmid and be packaged like a phage. They can insert up to 50kb.
4. **Yeast artificial chromosomes (YAC)**- primarily used in genome sequencing projects. They host large inserts up to 1000kb.

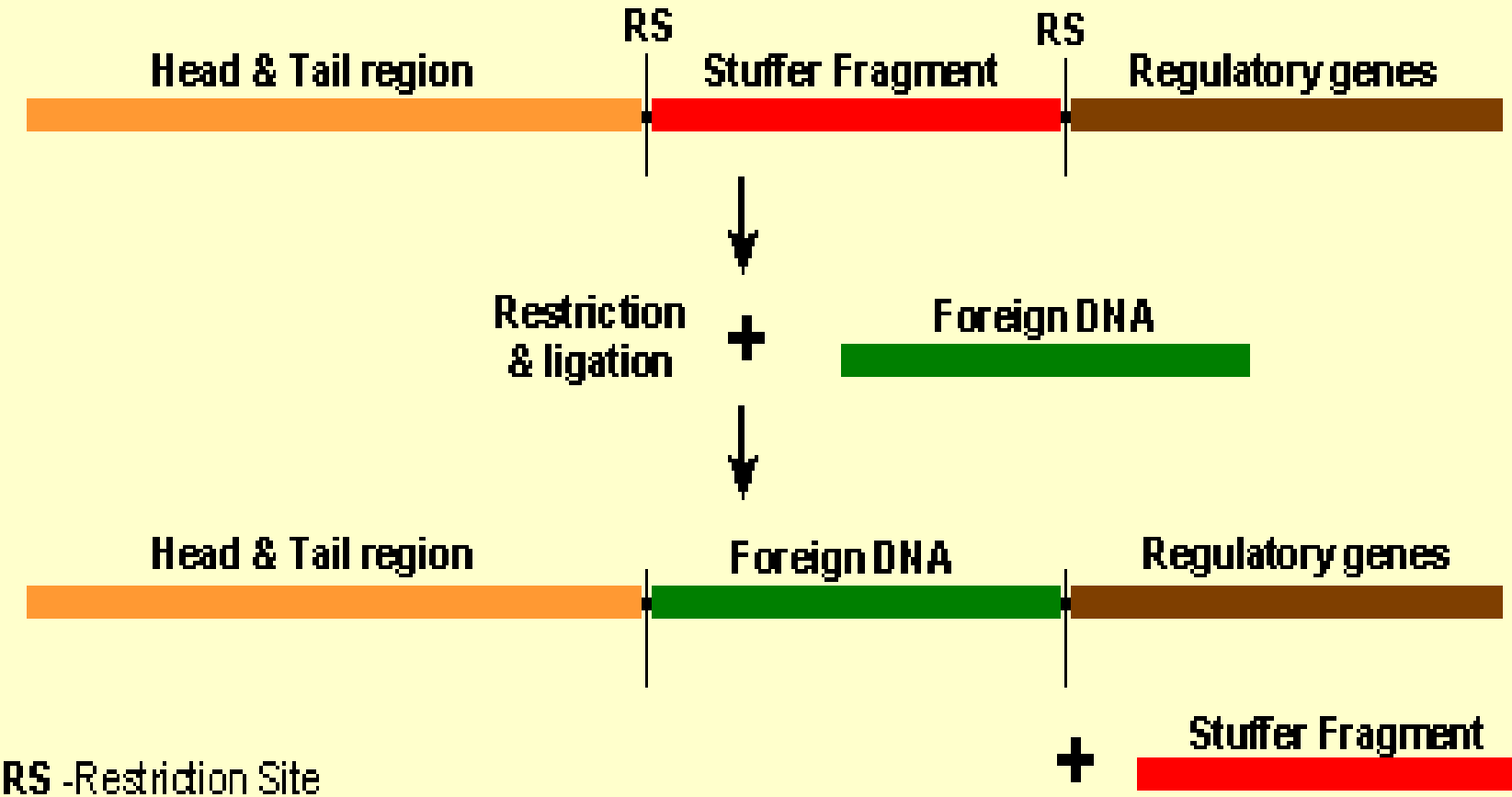
Vector	**Insert size (kb)**
Plasmid	<10 kb
Bacteriophage λ	9-15 kb
Cosmids	33-50 kb
YACS	100-1000 kb

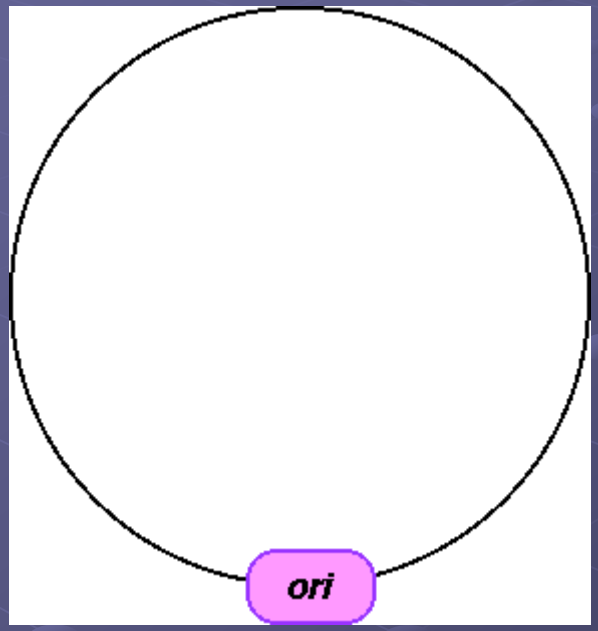
What determines choice of vector?

1. Insert size
2. Vector size
3. Restriction sites
4. Cloning efficiency

-Central 1/3 is the “stuffer” fragment.
-Segments of the lambda DNA are removed and a stuffer fragment is put in, this keeps the vector at a correct size

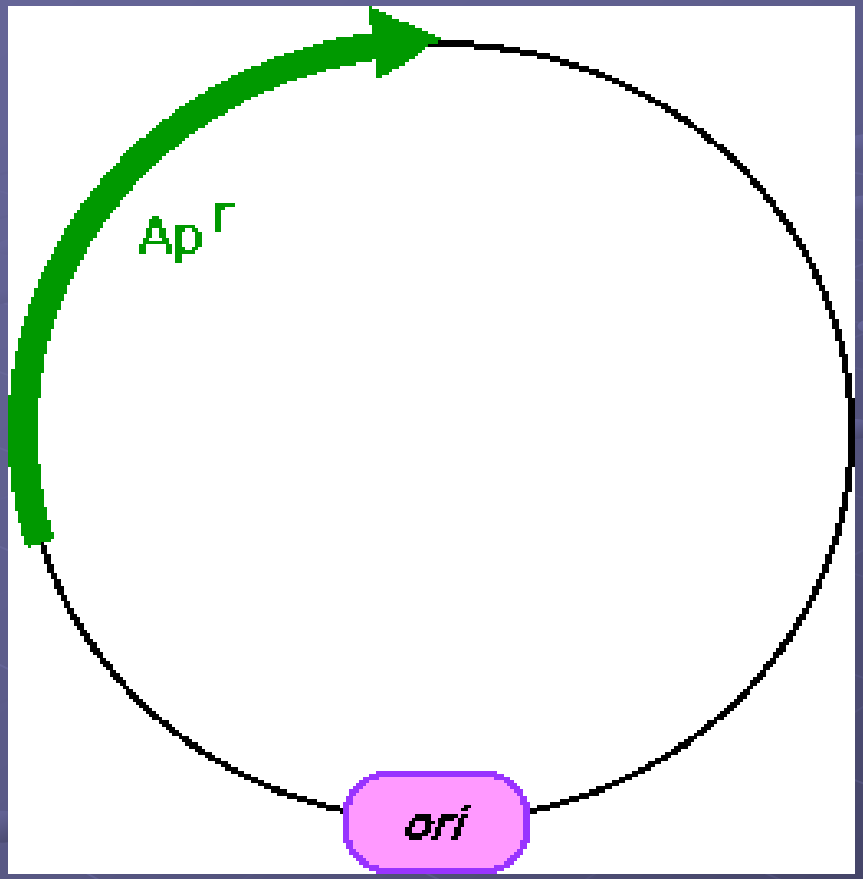
Lambda Replacement Vectors





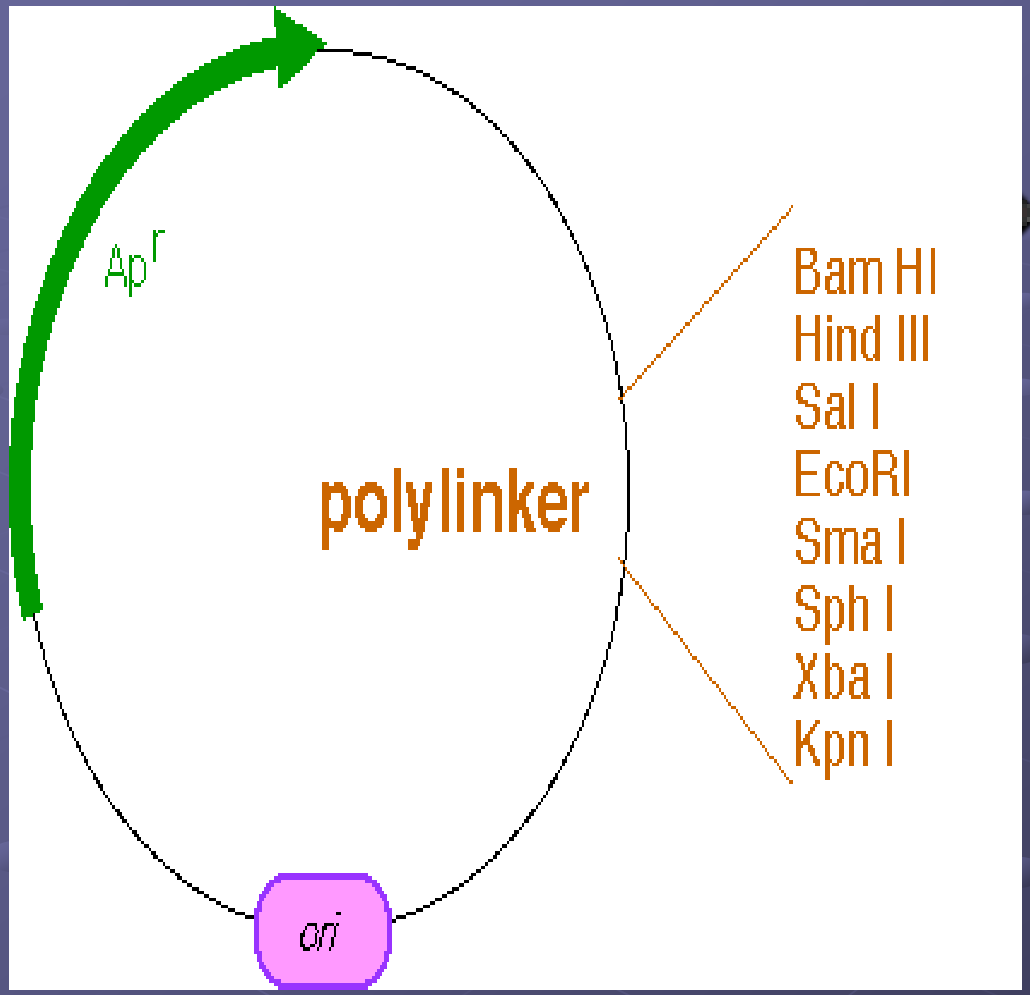
- **Origin of replication** is a DNA segment recognized by the cellular DNA-replication enzymes. Without replication origin, DNA cannot be replicated in the cell

http://www.uic.edu/classes/phar/phar331/lecture6/



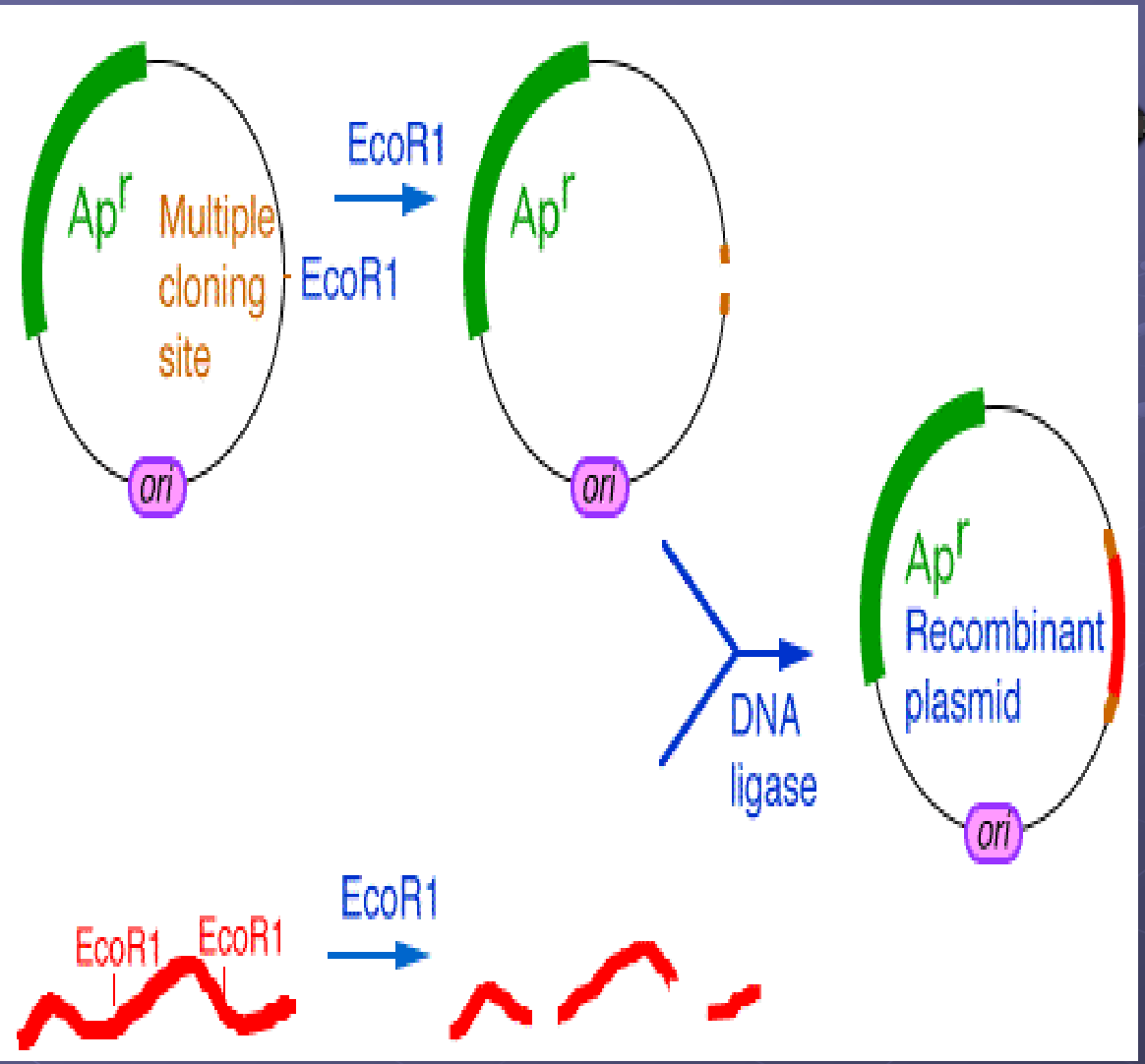
- **Selective marker** is required for maintenance of plasmid in the cell. Because of the presence of the selective marker the plasmid becomes useful for the cell. Under the selective conditions, only cells that contain plasmids with selectable marker can survive.

http://www.uic.edu/classes/phar/phar331/lecture6



- Many cloning vectors contain a **multiple cloning site (**DNA segment with several unique sites for restriction nucleases located next to each other)

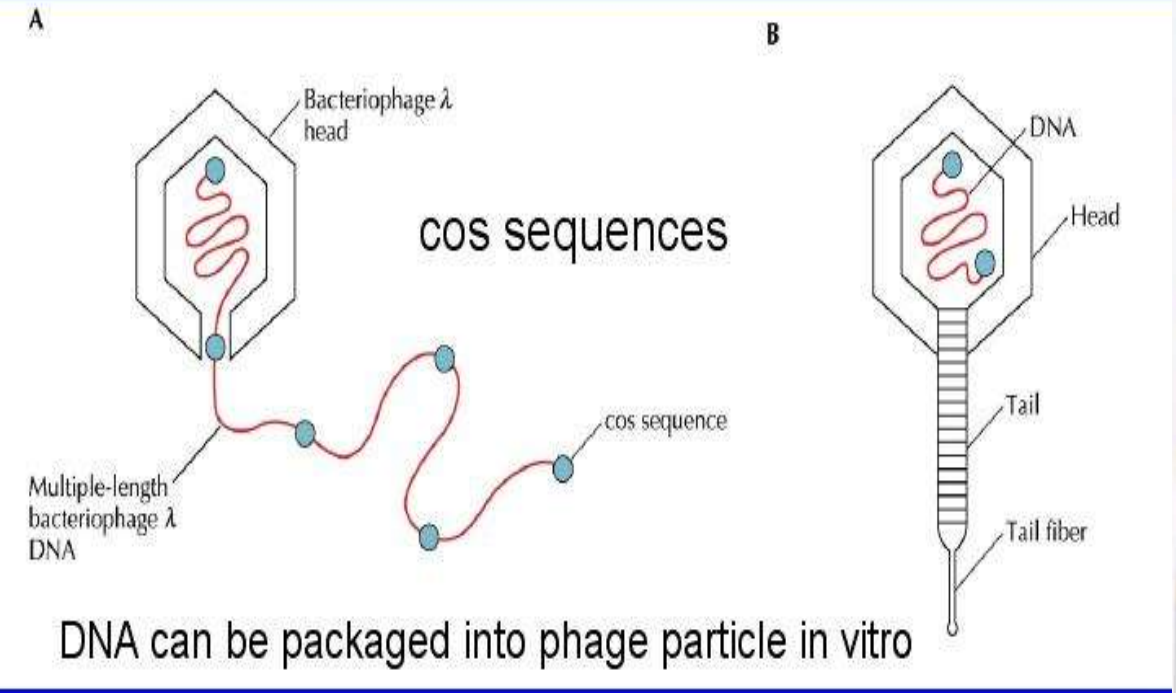
http://www.uic.edu/classes/phar/phar331/lecture6



- Gene to be cloned can be introduced into the cloning vector at one of the restriction sites present in the cloning site.

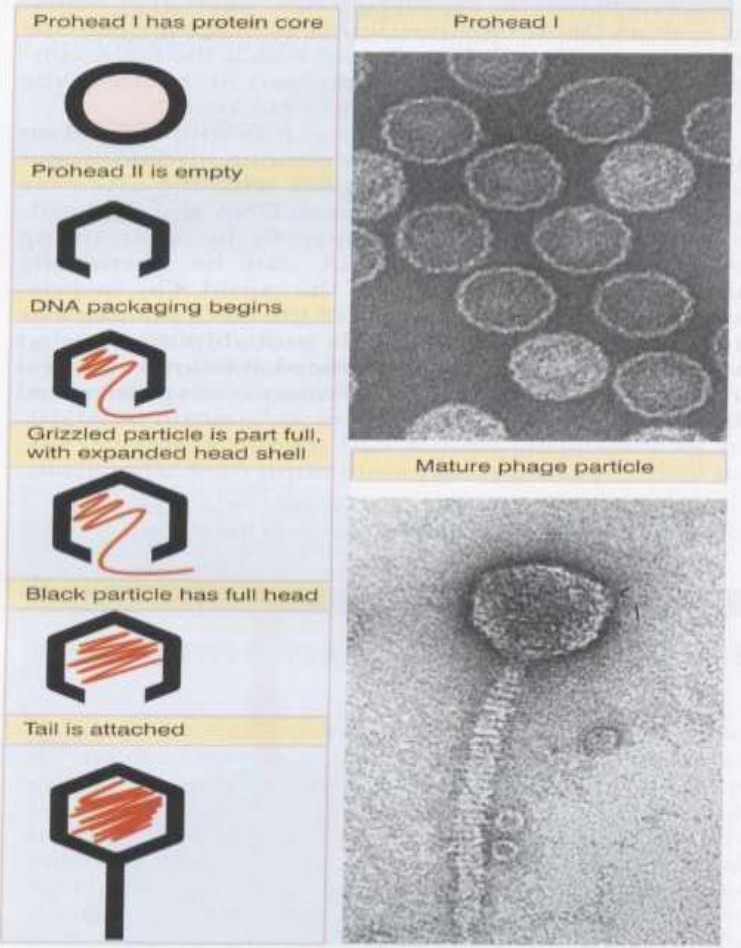
http://www.uic.edu/classes/phar/phar331/lecture6

Fig 4.24 Bacteriophage Lambda (λ)



DNA can be packaged into phage particle in vitro

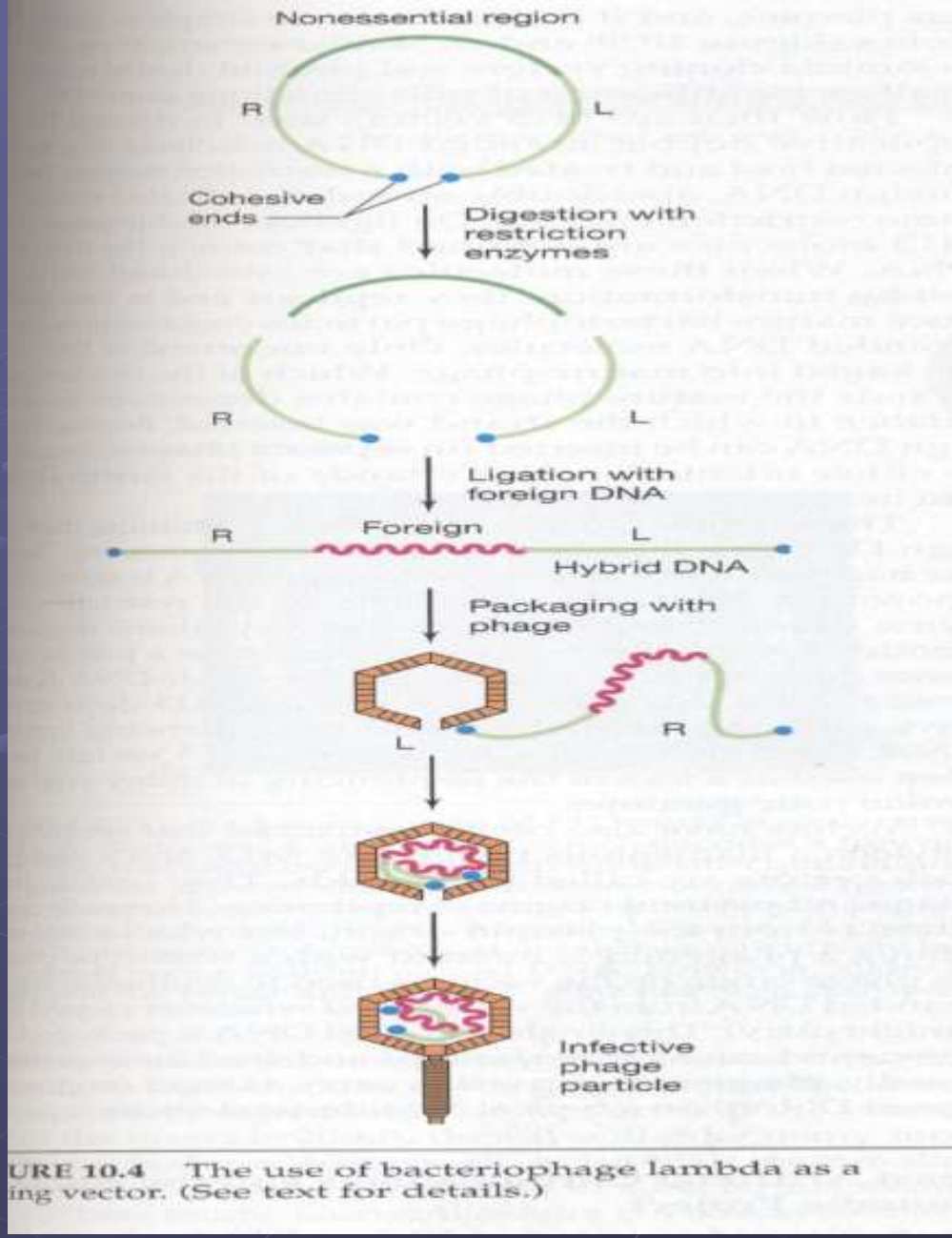
Figure 18.3 Maturation of phage lambda passes through several stages. The empty head changes shape and expands when it becomes filled with DNA. The electron micrographs show the particles at the start and end of the maturation pathway. Photographs kindly provided by A. F. Howatson.



http://www.gmu.edu/departments/biology/385-Ch04c-rDNA/

steps in cloning with λ :

- Isolate vector DNA and gene of interest
- Cut both with restriction enzyme(EcoRI)
- Connect two fragments of foreign DNA with DNA ligase. (recombinant DNA)
- Package DNA by adding cell extracts containing head and tail proteins
- Transfer recombinant molecules into host cell (transform)
- Grow/select transformants: check recombinant phage for the presence of desired foreign DNA sequence by observing its genetic properties.



URE 10.4 The use of bacteriophage lambda as a ing vector. (See text for details.)

Brock Biology of Microorganisms, 9th Edition (2000)
Prentice Hall, Madigan, Martinko, Parker

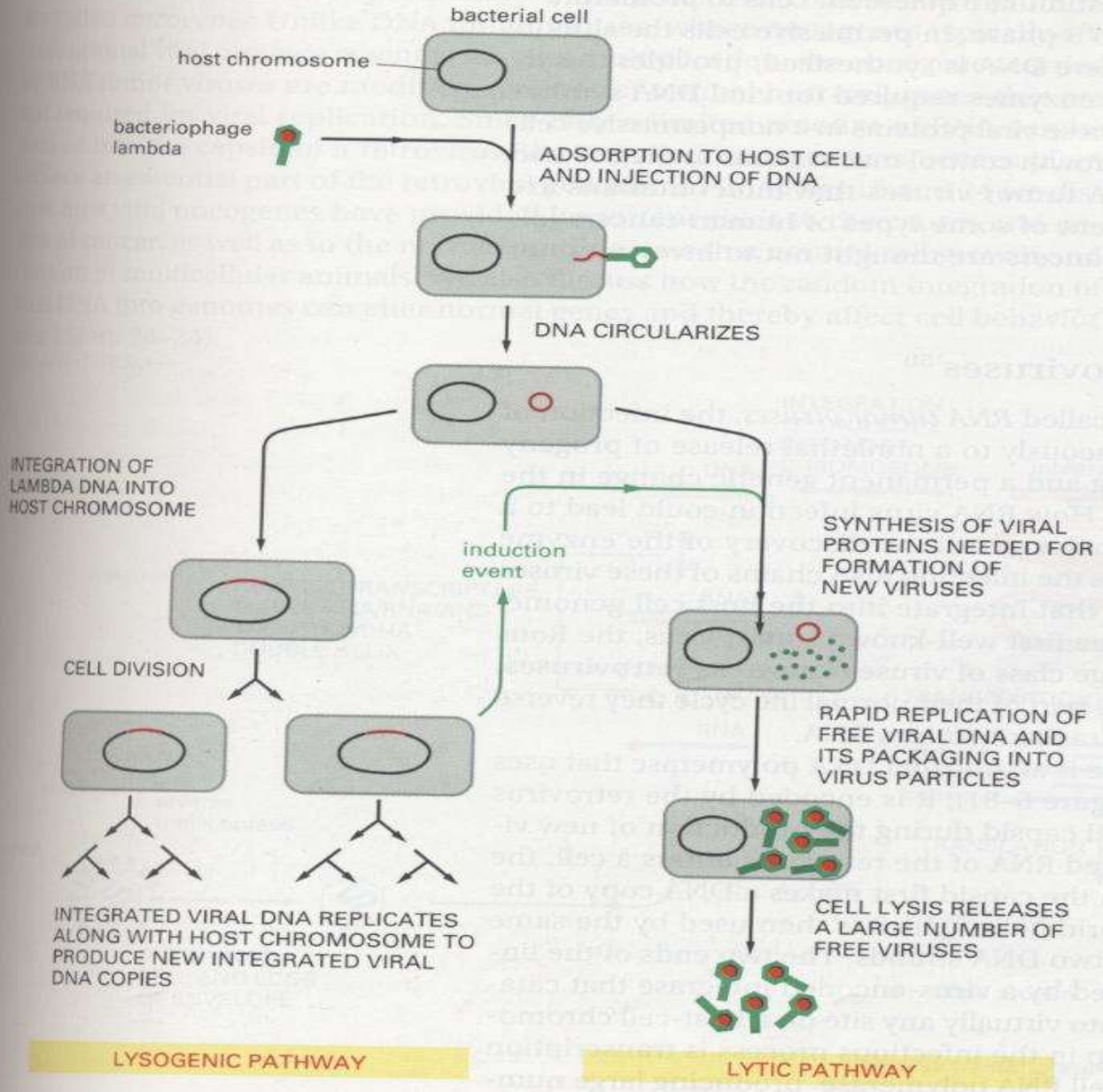


Figure 6–80 The life cycle of bacteriophage lambda. The lambda genome contains about 50,000 nucleotide pairs and encodes about 50 proteins. Its double-stranded DNA can exist in either linear or circular forms. As shown, the bacteriophage can multiply by either a lytic or a lysogenic pathway in the *E. coli* bacterium. When the bacteriophage is growing in the lysogenic state, damage to the cell causes the integrated viral DNA (provirus) to exit from the host chromosome and shift to lytic growth. The entrance and exit of the DNA from the chromosome are site-specific genetic recombination events catalyzed by the lambda *integrase* protein (see Figure 6–68).

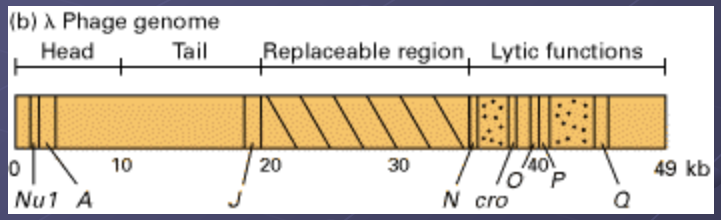
Molecular Biology of the Cell, 3rd Edition, Garland Publishing, Inc. 1983

λ-phage as a Vector

- The genomic library is generated by using λ-phage for the following reasons.

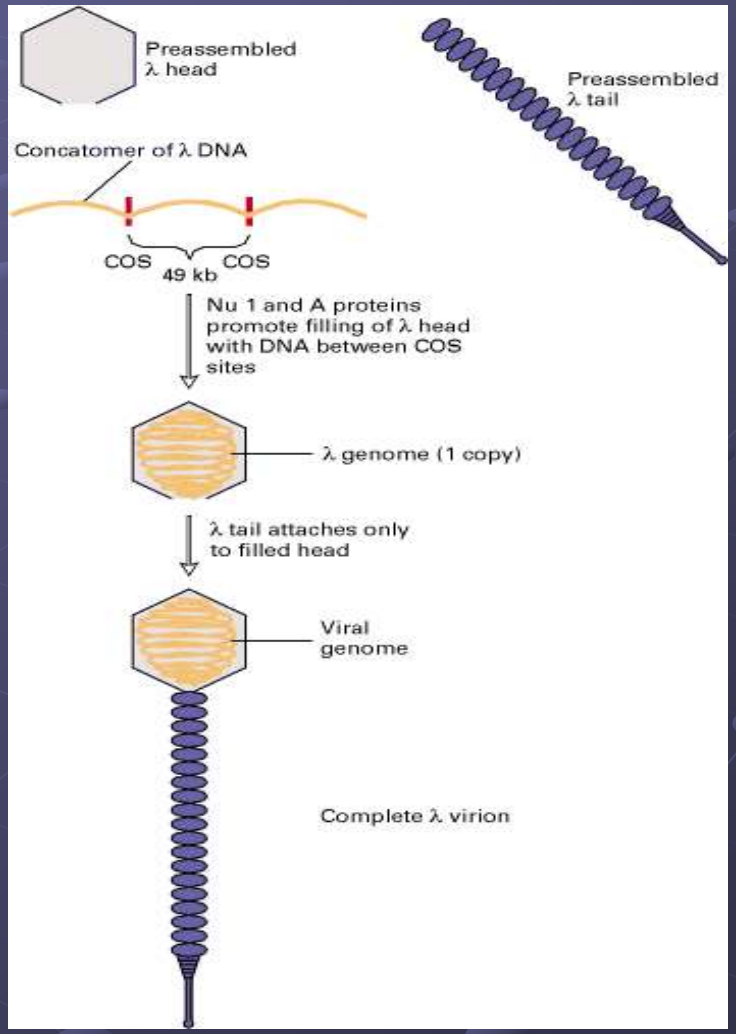
1. A large number of λ phage can be screened simultaneously (5 x 10E4 phage plagues).
2. λ phage as a higher transformation efficiency about 1000 times higher compared to a plasmid.

- The vector as to maintain its lytic growth.
- Lysogenic pathway and other viral genes that are not important are replaced with the DNA to be cloned.



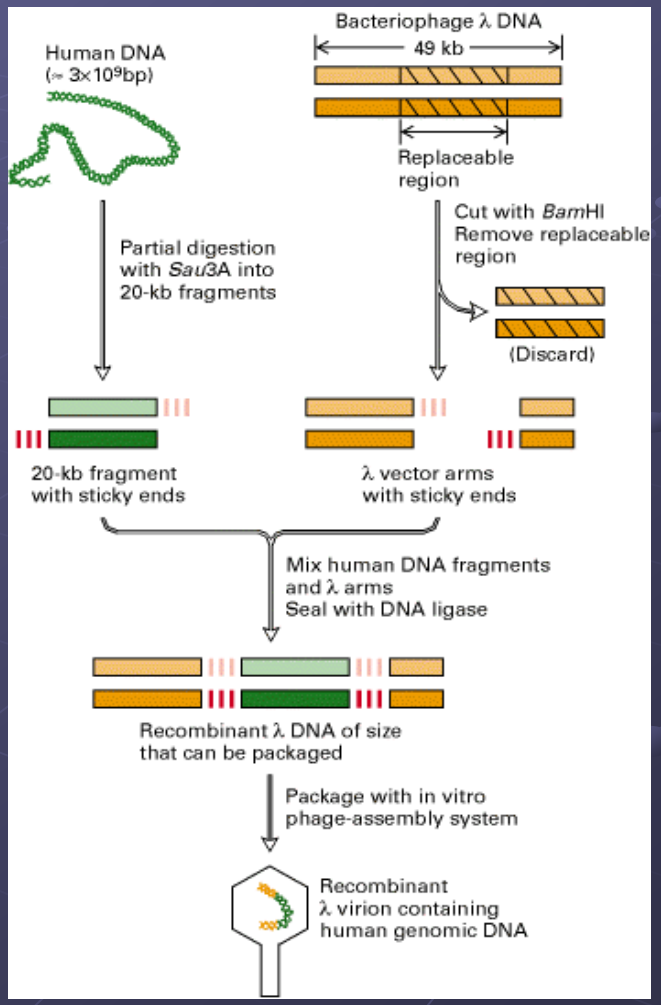
λ-phage as a Vector (Cont.)

- An infected E.Coli will produce what are know as *concatomers* (which is the viral genome) on either site of the *concatomers* there is a site called COS Site.
- Two proteins recognize this site A protein and Nu protein, which will lead to the insertion of the λ DNA into the phage head. The chromosomal DNA that lacks the COS sites will not enter the phage head. Once the genetic information is inserted the tail will assemble.
- A 50kb can be inserted into the phage.



Generating A Genomic Library

- λ-phage is treated with restriction enzymes that produce λ arms with sticky end. These arms contain all the lytic genetic information that is needed for replication and produces room for insertion of new genetic information.
- DNA sequence is obtain from the cell of interest. It is cleaved with restriction enzymes that produce 20kb fragments that have complementary sticky ends.
- Both are mixed in equal amounts and are treated with a DNA ligase that cleaves them together.
- Afterward the entire combined sequence is packed to the phage head.

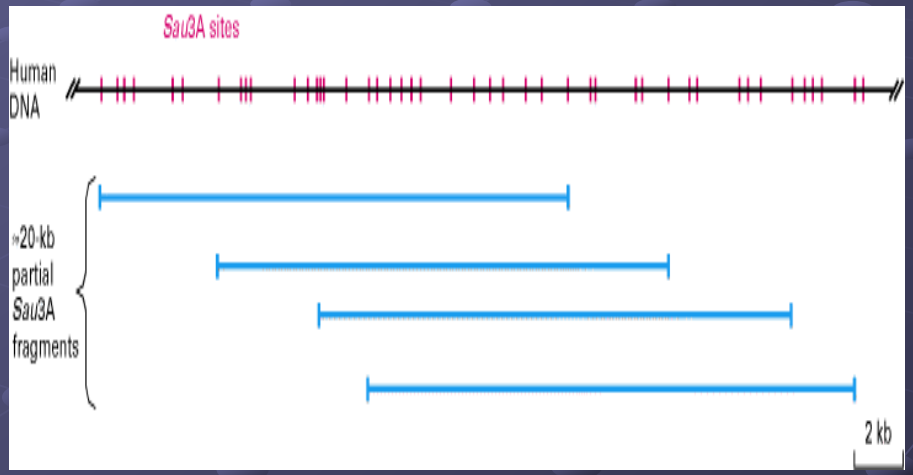


Packaging of the Recombinant DNA

- To prepare the phage an E.coli cell is infected with a mutant λ-phage that as a defective “A-protein” (which is one of two genes that are responsible for packaging genetic information).
- Therefore the E.Coli accumulates empty heads and also preassembled tails.
- Once enough heads and tails are assembled we lysate the E.Coli cells.
- To the mixture of heads and tail we add isolated A protein (obtained from E.Coli infected with λ-phage).
- In the next step we add the recombinant DNA that has the λ-phage genetic information (which also includes COS sites).
- At this point we have a mixture containing mutant λ-phage heads and tails. There is isolated A protein and recombinant DNA containing λ-phage genetic information with COS sites.
- Therefore we have all the components necessary to package the recombinant DNA into the λ-phage head. Once the information is inserted the tail assembles and we have an infectious phage that contains the recombinant DNA sequence.

Generating A Genomic Library (cont.)

- In order to sequence the entire genome it is necessary to produce overlapping sequences.
- Using a technique called chromosome walking, it is possible to order to genomic clones.
- As I mentioned earlier the human genome contains 3 × 10E9 base pairs. Each vector contains 20kb that means that it is necessary to generate about 1.5 x 10E5 phages.
- You can screen 5 × 10E4 plaques on each petri dish meaning that you can contain all the human genome on 20-30 petri dishes.
- If plasmids were used instead of λ phage it would take 5000 petri dishes.

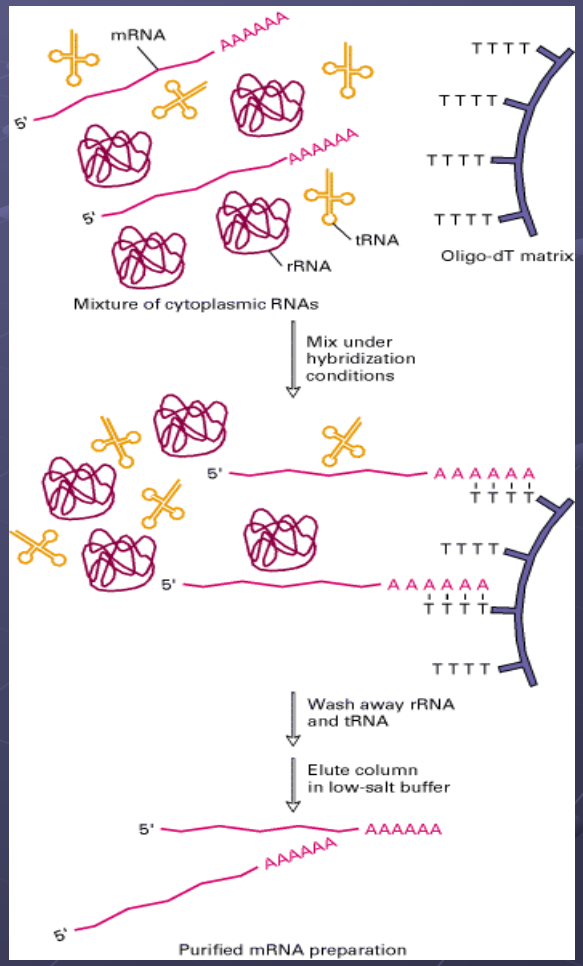


Generating A cDNA Library

- All eukaryotes have an mRNA. Each specialized cell have a specific mRNA that encodes information for a specific protein.
- This information can be transformed back into DNA or in other words a DNA copy of mRNA (using reverse transcriptase). This cDNA can be stored in plasmids or phages.
- cDNA contains only the expressed genetic information which allows us to study the amino acid sequence directly from the DNA.

Generating A cDNA Library(Cont.)

- The first step in creating a cDNA library is to isolate mRNA from the cell.
- All mRNA have a poly A tail (unlike tRNA and rRNA that don't). By using a column that contains a short poly T sequence it is possible to isolate the mRNA for both tRNA and rRNA.

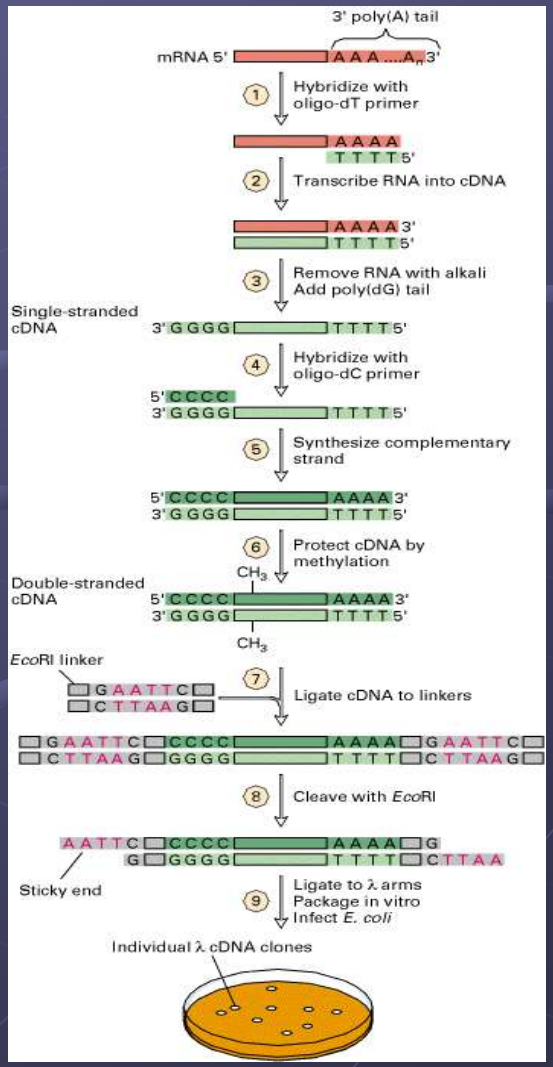


Generating A cDNA Library(Cont.)

- Once the mRNA is isolated it is treated with an enzyme reverse transcriptase (which is found in retro viruses like HIV). This enzyme will create a (ss)cDNA intermediate from the mRNA.
- By hybridizing the poly A of the mRNA with oligo T'd a primer is created. Reverse transcriptase recognizes this template and will add bases to 3' end.

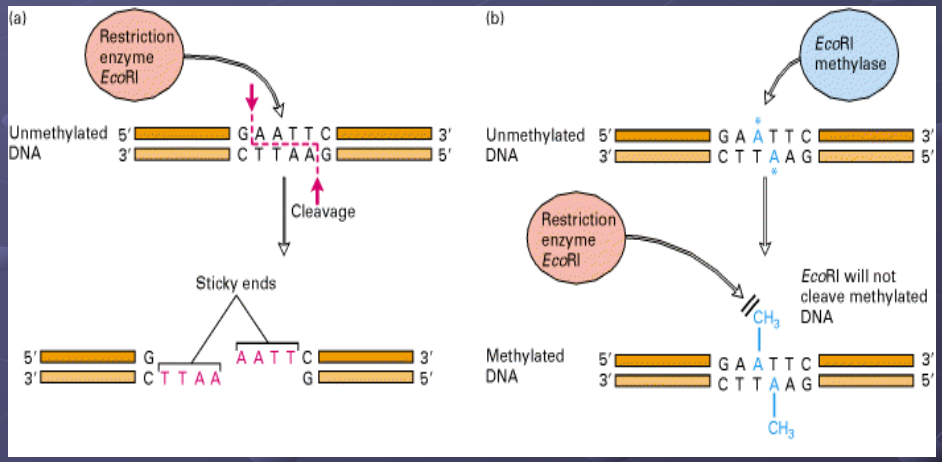
Generating A cDNA Library(Cont.)

- At this point the (ss)cDNA needs to be converted to the double strand cDNA.
- The mRNA cDNA complex is treated with an alkali which hydrolyzes the mRNA, but not the cDNA.
- Then by using *terminal transferase* which is a DNA polymerase that adds deoxynucleotides to free 3 ends without the need of template (this will add poly G).
- To this a synthetic poly C is hybridized which is used as primer for the synthesis of the complementary strand of the cDNA.



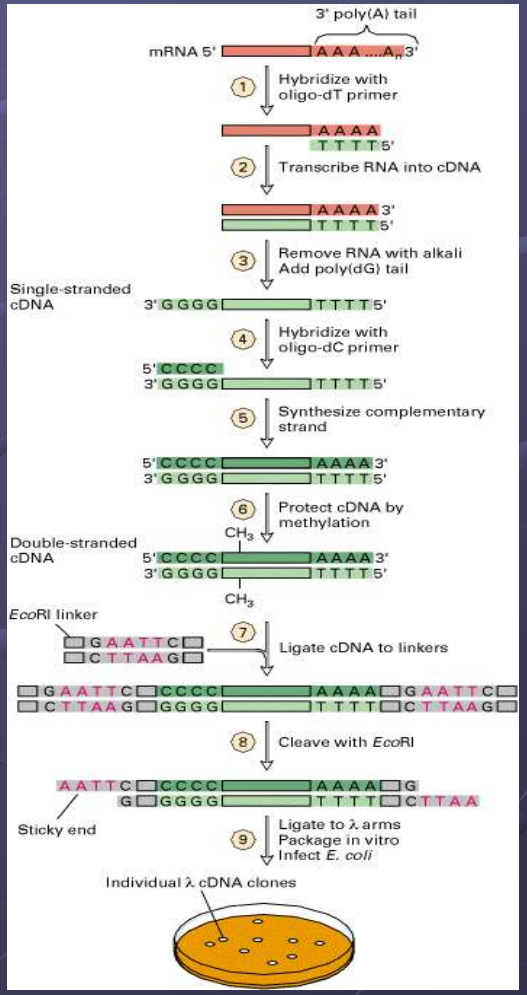
Packing the cDNA

- The first step is to ligate to each end of the cDNA a short *restriction-site linkers* (which are prepared by bactiophage T4). This will produce blunt end at the end of the DNA.
- Now it is necessary to protect the cDNA from unwanted digestion by restriction enzymes. Therefore the cDNA is treated with a modification enzyme that methylates specific bases within the restriction enzyme sequence.
- The next step is to treat the cDNA with restriction enzymes that are specific to the blunt ends. This will result with sticky end.



Packing the cDNA(cont.)

- The final step is to ligate the sticky ends of the cDNA with the λ-phage arms that have complementary sticky ends, thereby inserting the Double strand cDNA into the vector.



References

- http://www.ncbi.nlm.nih.gov/books/bv.fcgi?call=bv.View..ShowSection&rid=mcb.section.1611
- **Molecular Cloning:** A Laboratory Manual Joseph Sambrook, David W. Russell