BIOC4004 - Industrial Biochemistry
Lecture 03 - Mon Jan 12, 04

Topics for the Day:

The Basics of recombinant DNA technology

Methods in Recombinant DNA technology

- Basic Enzymatic manipulation of DNA
- Cloning, Cloning vectors, Cloning strategies
- Southern, Western, Northern blots
- The Polymerase Chain Reaction (PCR)
    - Novel applications of PCR
- DNA sequencing
    - manual vs automated

Oral Presentation Topics:

- Improvement of industrial enzymes through protein engineering
- Pathway engineering to produce useful compounds
- Harnessing microorganisms for Bioremediation
- Development of novel antibiotics
- Genomes of exotic species; gene prospecting
- Genomics/Proteomics approaches to drug design
- Novel diagnostic techniques (forensics, medicine, GMOs)
- Engineering in Agrifood sciences
- Vaccine or Antibody engineering
- Applications of microarray technology
- Mining genome data for Industrial purposes
- Novel Bioinformatics tools

- note 1: if you'd like to propose a new topic, be my guest but please get my approval
- note 2: find yourself three partners in crime
- note 3: pick a topic
- note 4: each group submits three choices in order of preference **(by Thursday, Jan 15 by e-mail: ed.taboada@nrc.ca)**
- note 5: first come, first served
- note 6: I will announce each group and topic on Friday January 16th

Recombinant DNA (rDNA) Technology

A set of techniques used to :

- Isolate, manipulate, modify,mass produce DNA sequence of interest
- Analyze sequence of interest

The Problems:

- sequences of interest are interspersed with thousands of other sequences
- typically we're interested in genes that are found in 1 copy per genome
- we will need **lots of copies** for analytical and manipulation purposes

The Solutions:

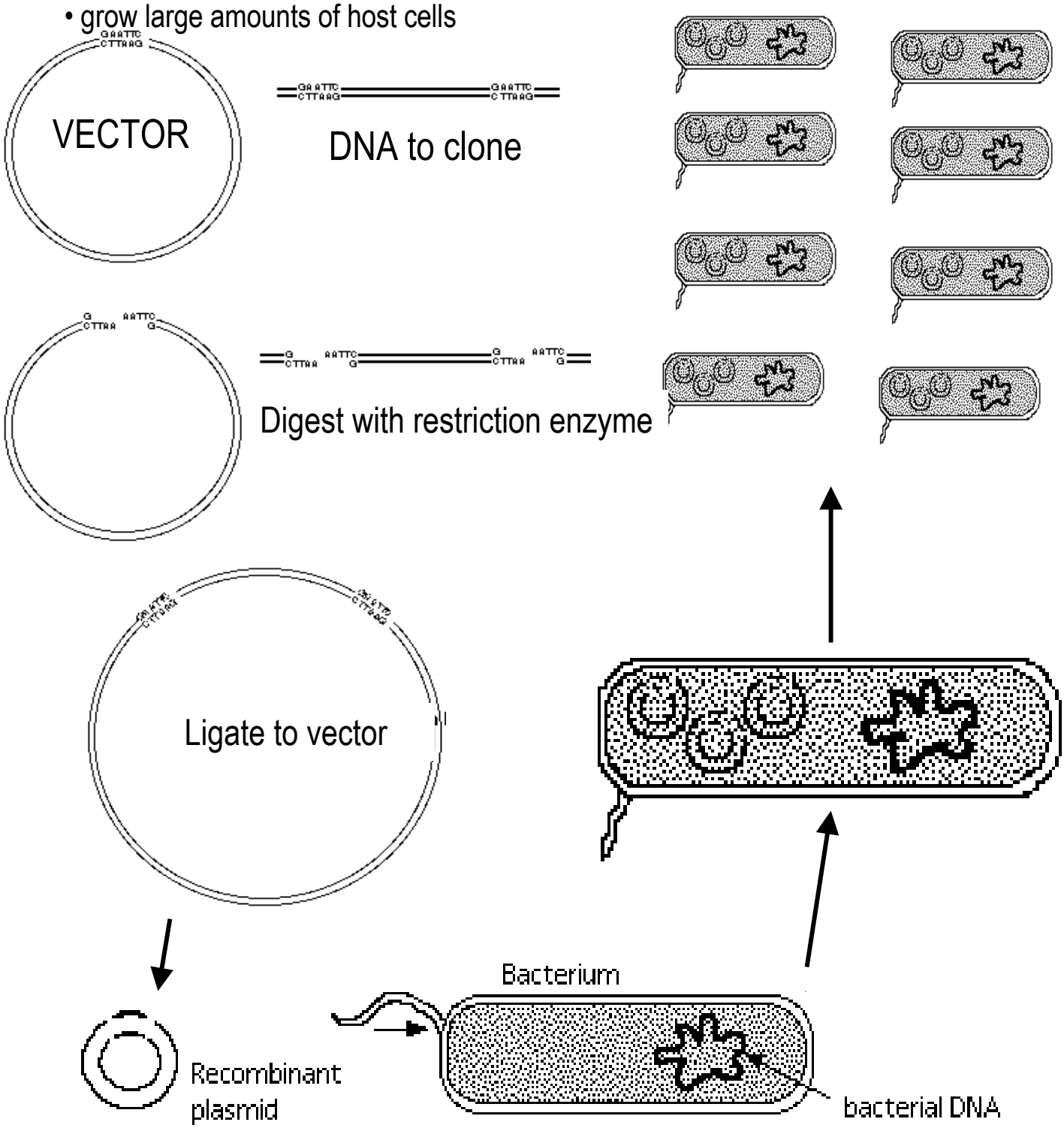
- Techniques for detection of specific sequences in complex mixtures
- Techniques for isolation and enrichment of specific nucleic acids
- Techniques for the manipulation of DNA sequences
- Techniques for large-scale amplification of a given sequence
- Analytical techniques for nucleic acids

The Engine of Recombinant DNA Technology is CLONING

- Generating **clones**
  - large number of cells identical to each other, derived from a single initial cell
- Clones are carrying specific DNA sequences of interest
- A **vector** is used to carry the DNA of interest within each clone
- Process involves:
  - joining pieces of DNA from different sources to yield a novel molecule (ie. a **recombinant** molecule)
  - introducing vector into a cell
  - growth and division of the cell
    - ie. amplification of the DNA contained within cell

Generalized Scheme for Cloning

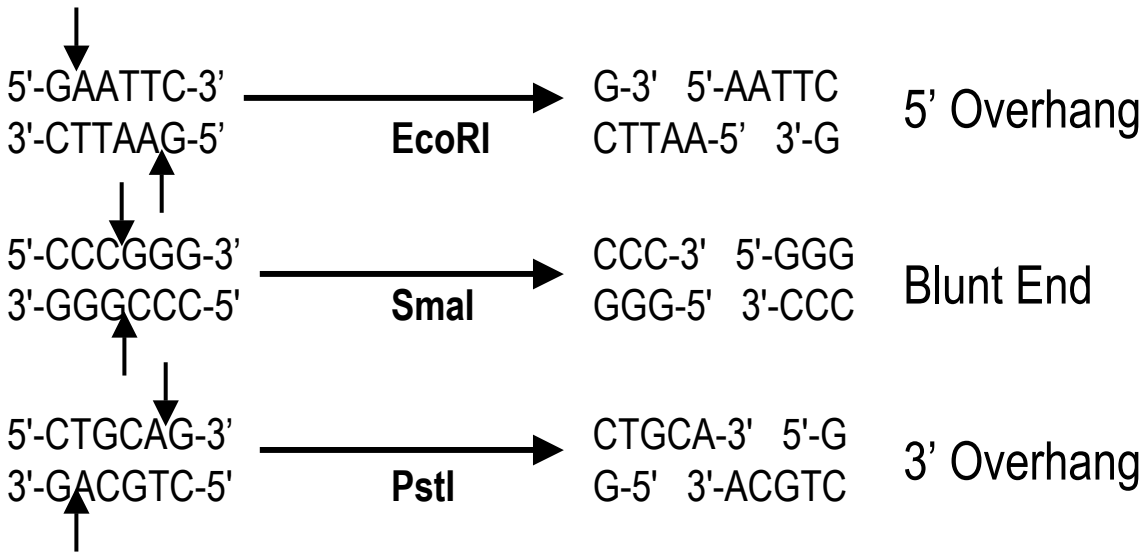
- paste gene of interest into vector
- put vector into bacterial cells (or other host cell)
- grow large amounts of host cells



Type II Resctriction Endonucleases

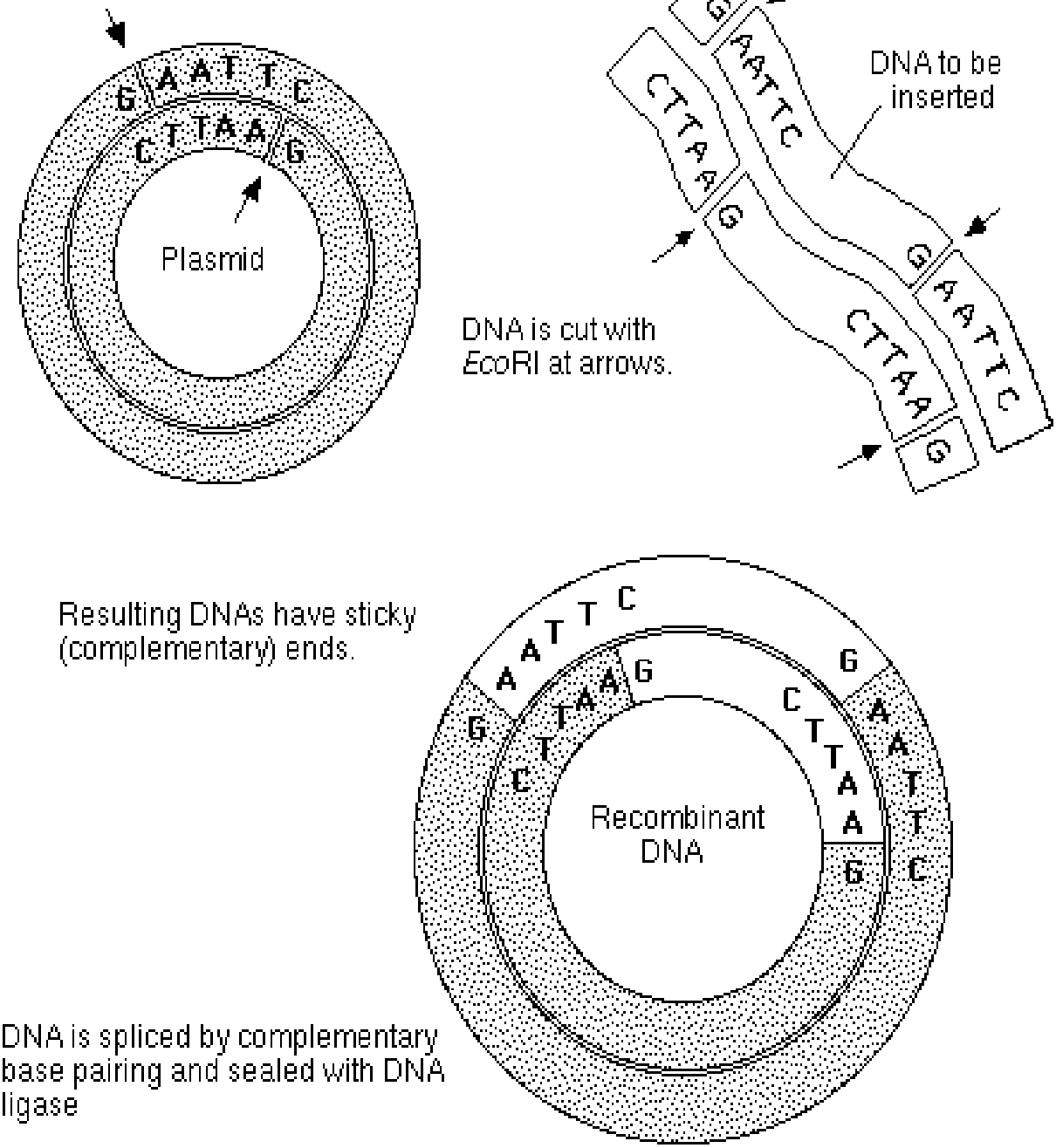
- Recognize specific sequences and cut the DNA at the recognition site
- Recognition sequence can be 4-12 bp long
  - the longer the recognition site, the less frequently it will be found
  - base composition has direct effect on frequency of recognition sites
- **Source:** bacteria (they use them to destroy foreign DNA)

• **EcoRI**	Escherichia coli	GAATTC
• **SmaI**	Serratia marcescens	CCCGGG
• **PstI**	Providencia stuartii	CTGCAG



- recognition sites are typically palindromic
- 3' and 5' overhangs are called "Sticky Ends"
- Number of bases in overhang is enzyme-specific
- *Isoschizomers*: same recognition sequence, same or diff. sticky ends

Restriction Enzymes can be used to generate rDNA



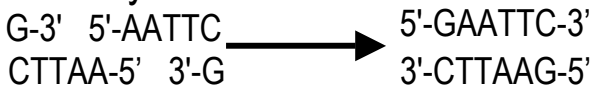
How do we join two fragments together ? **DNA ligases**

DNA Ligases

- bacterial or bacteriophage origin

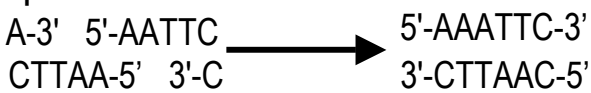
- "seal-up" nicks in DNA
  - reforms phosphodiester bond between adjacent bases

- Can join:
  - sticky ends
    - same enzyme or isoschizomers



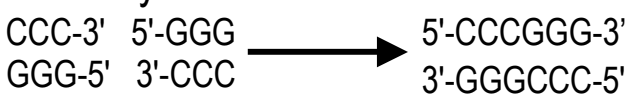
Recognition Sites Restored

    - compatible ends



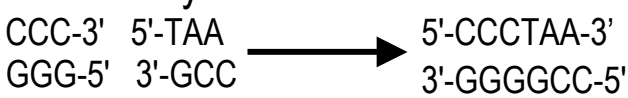
Recognition Sites Destroyed

  - blunt ends
    - same enzyme



Recognition Sites Restored

    - different enzymes



Recognition Sites Destroyed

Can join different things but you are not always able to re-cut!!!

Cloning and Vectors

Vectors are used to carry DNA of interest into host cells that can be grown in culture, thus amplifying the vector, and the sequences contained within.

Most Vectors feature:

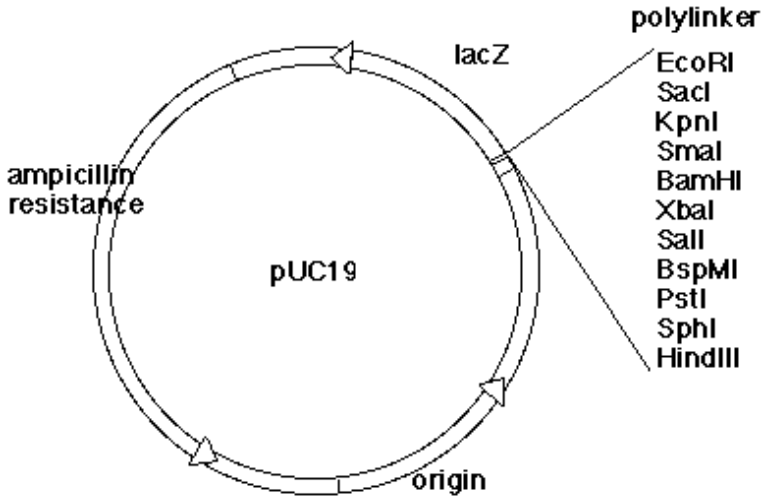
- Sequences for self-replication
- Selectable marker
- Screening marker
- Multiple Cloning Site
- promoter to transcribe cloned piece of DNA

Major Types of Vectors

- plasmid-based
- bacteriophage-based
- artificial chromosome-based

Plasmids: pBR322, pBluescript, pUC19….

- usually small, circular, double-stranded DNA molecules
    - <10kb (kb = kilobase, 1000 nts.)
- number of copies / cell varies
    - 1-2 copies: `**stringent**' or '**low-copy'**
    - many copies: `**relaxed**' or **'high-copy'**
- Selectable markers can be:
    - antibiotic resistance (eg. Ampicillin resistance)
    - auxotrophic selection (eg. Leucine biosynthesis)



- **shuttle-vectors**
- **single-stranded / double-stranded** switching in some

Advantages

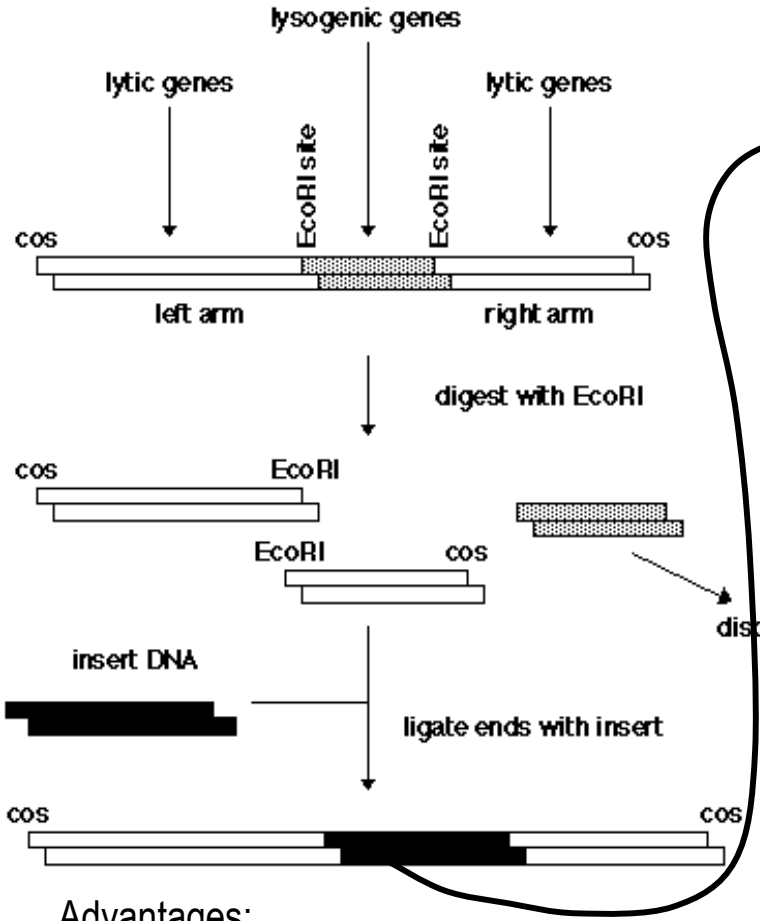
- small (easy to work with)
- easy to purify
- Poly-linker gives many restriction enzyme options

Plasmid disadvantages

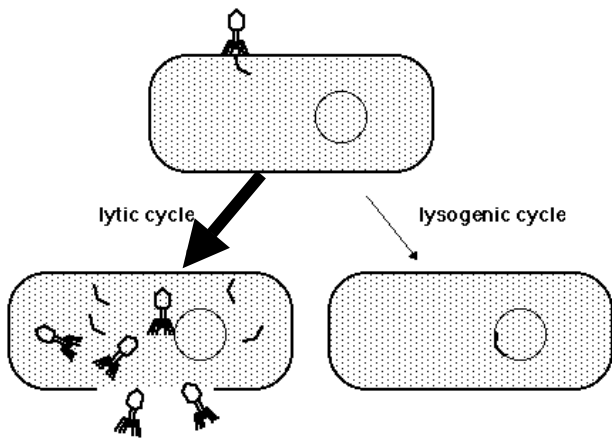
- relatively low transformation efficiency
- only relatively small inserts allowed (100 bp to ~12-15 Kb)

Phage-based Vectors: eg. Lambda-based vectors

- Based on *E.coli* bacteriophage Lambda (a virus)
- Wild-type size is 48,502 bp and sequence is known
- A chunk in the middle is dispensable (20-30 Kb)
  - replace with DNA of interest (can also insert small DNA < 10 Kb)
  - efficient phage "packaging": total 38 to 52 Kb



- recombinant Lambda can be "packaged" into phage particles
- transfect *E.coli*
- Lytic cycle produces LOTS of new phage



Advantages:

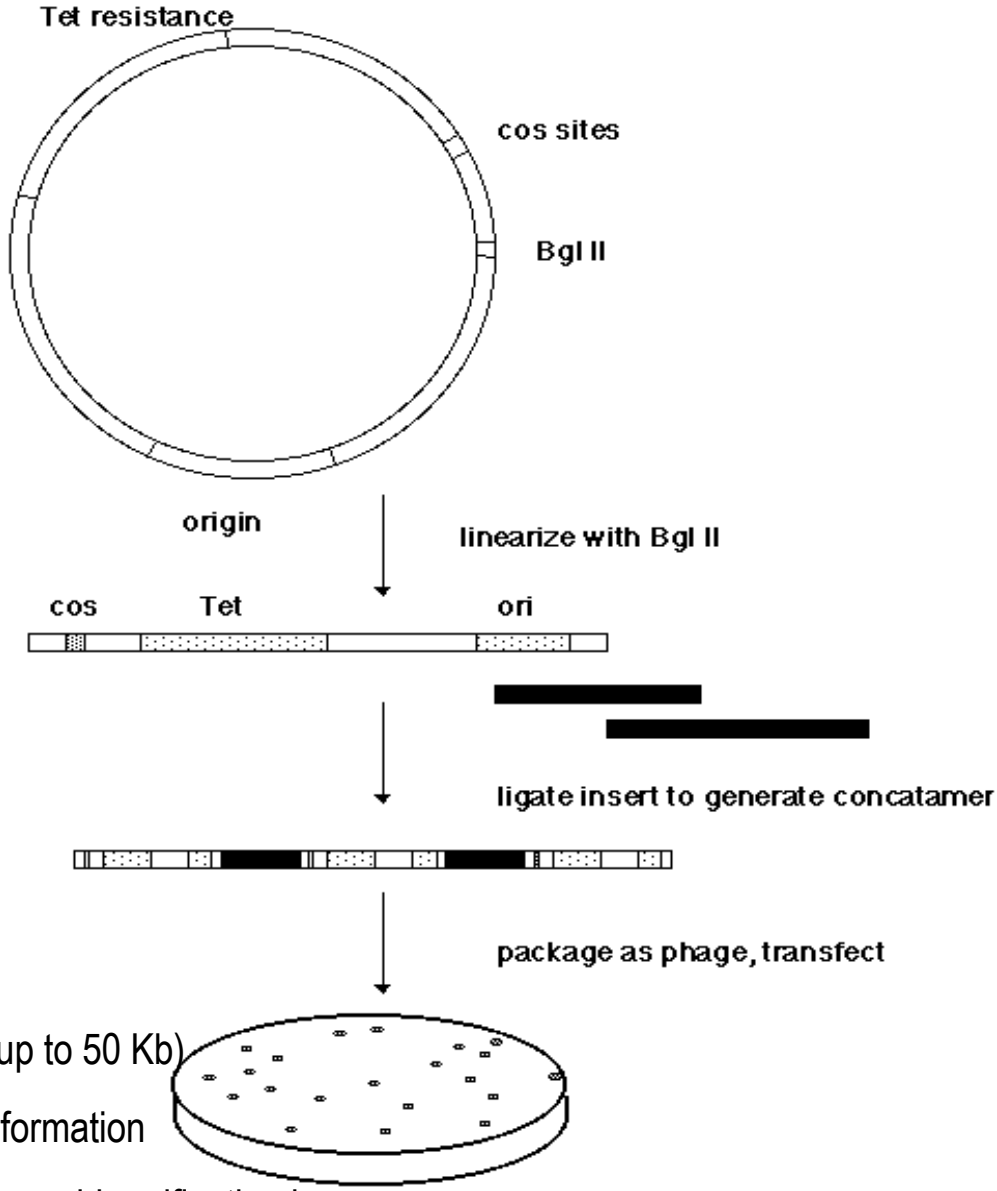
- accommodates bigger inserts than a plasmid (20-30 Kb)
- high efficiency of transfection

Disadvantages:

- limited number of restriction sites
- insert size is constrained (must have efficient packaging)
- have to package the DNA
- isolation of the DNA is more complicated than plasmids

A Phage / Plasmid Hybrid ? Cosmid vectors

- Engineered to act as large plasmids
- Handled as a plasmid (purification and manipulation)
- Engineered to package as a phage particle (Cos sites)
  - high efficiency of transfection



Advantages

- larger insert size (up to 50 Kb)
- high efficient transformation
- manipulation and cosmid purification is easy

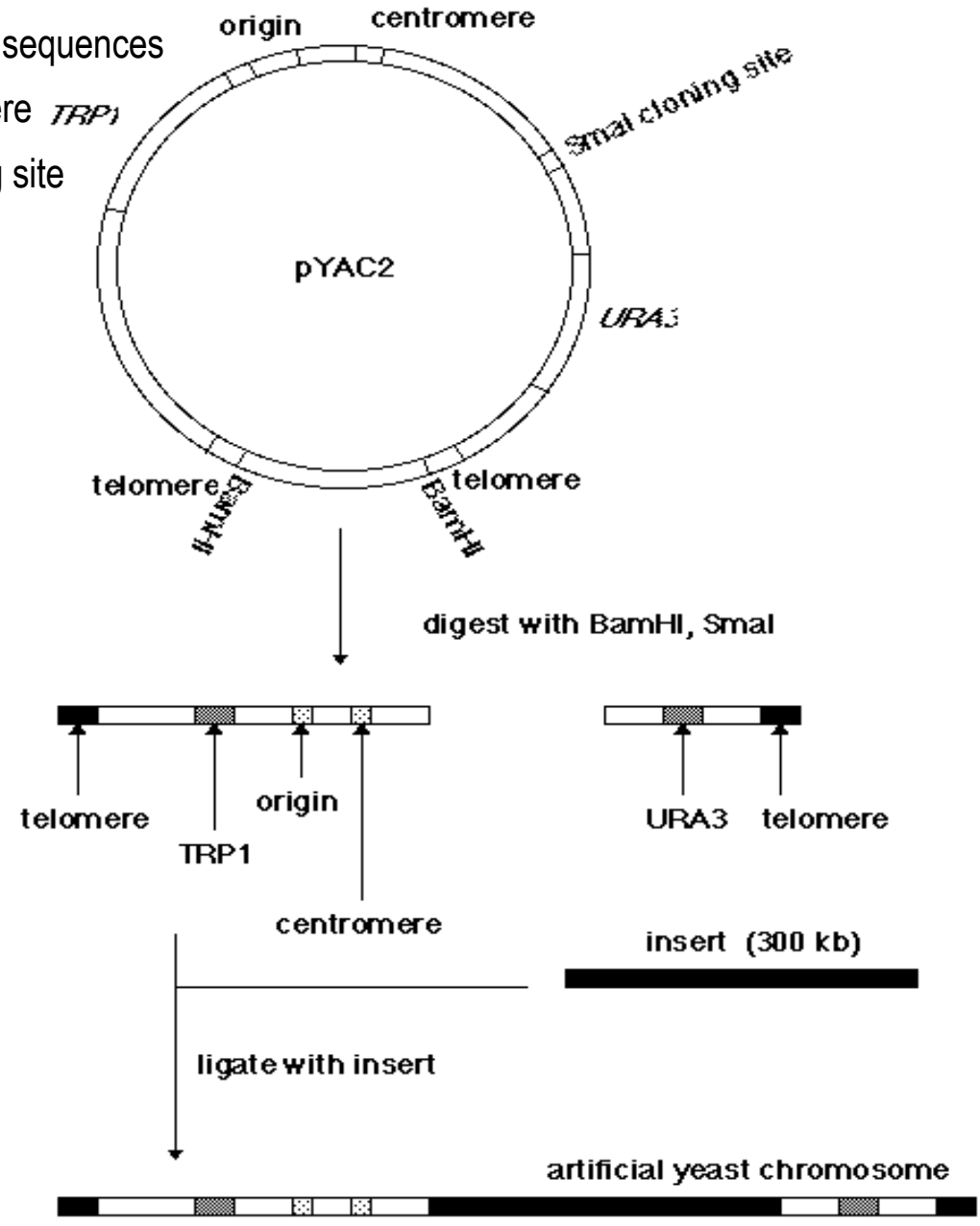
Disadvantages

- packaging is involved

Articificial Chromosome-based: eg. YACs (yeast articifial chromosomes)

• A plasmid with:

- yeast telomere sequences
- yeast centromere *TRP1*
- multiple cloning site



Advantages

- huge insert size (~300 Kb)

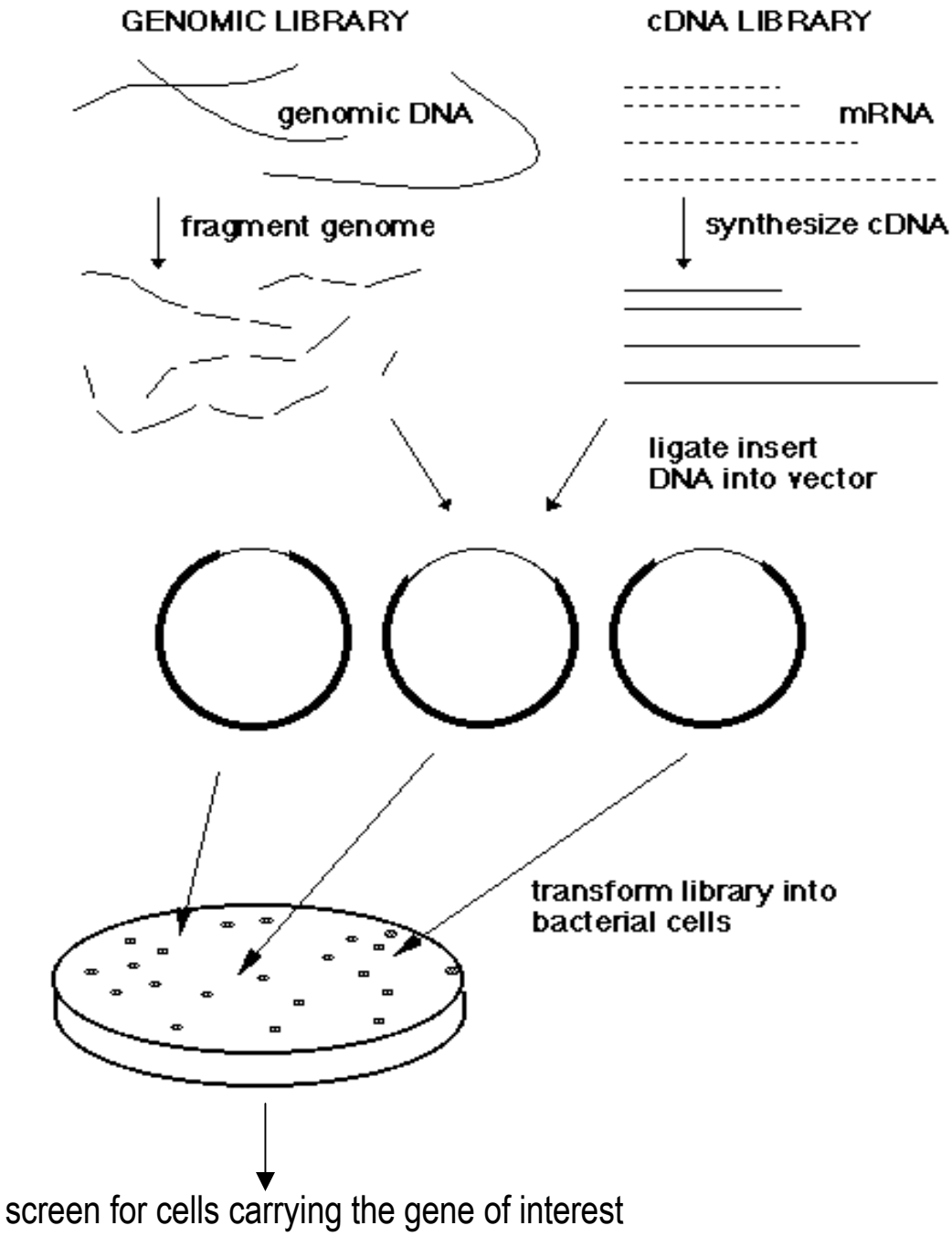
Disadvantages

- huge insert size can make DNA unstable (rearrangements)
- yeast host, not *E. coli;* much harder to work with

Recombinant DNA Libraries

• A collection of vector molecules carrying different inserts represent all or some of the genes from an organism or cell

• Used to find individual genes within complex genomes (**Screening**)

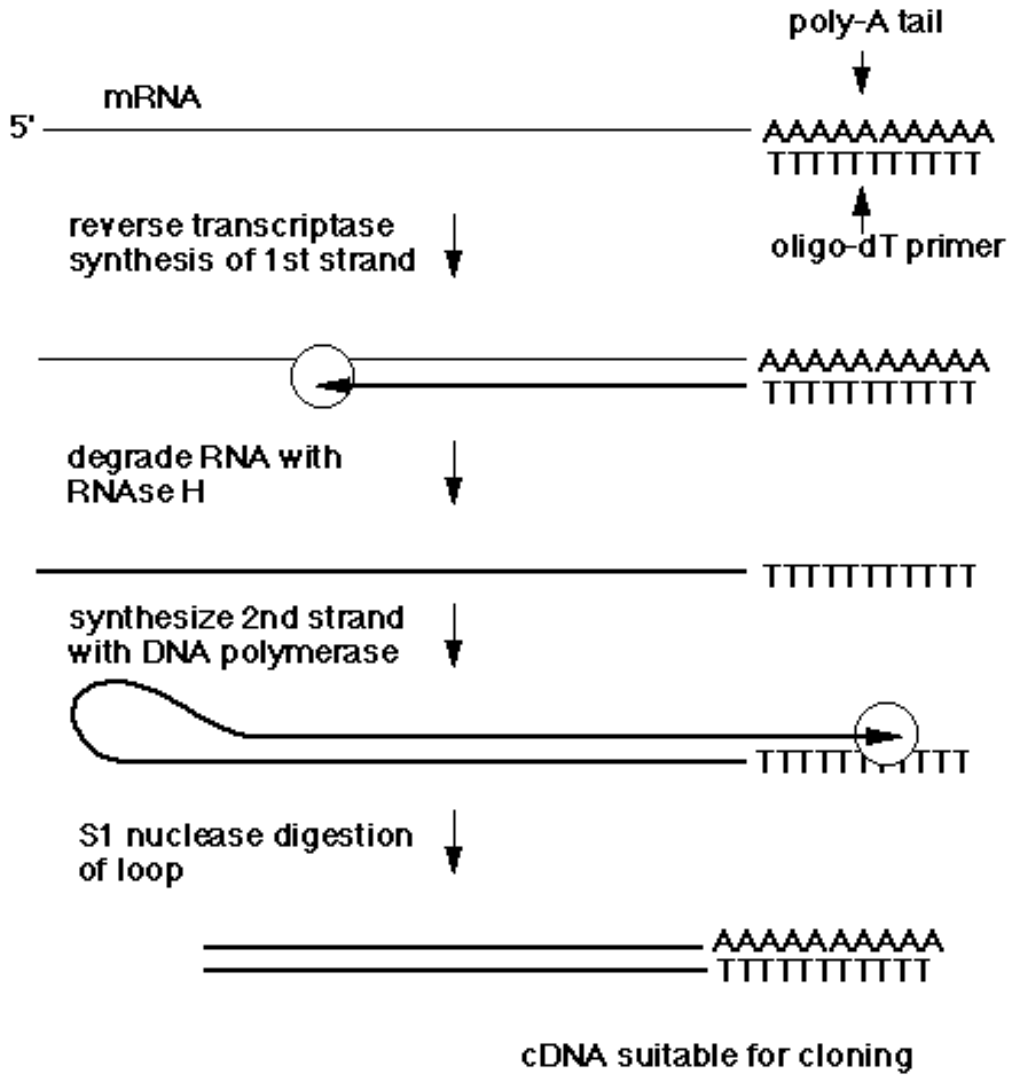


cDNA Synthesis

• Eukaryotic genes have introns, which are removed in mature mRNAs

• For most purposes (for example expression) the introns are a nuisance

• Solution: copy the mRNA into DNA (cDNA), which can be manipulated by standard recombinant DNA techniques



Cloning strategies: cDNA vs. genomic libraries

cDNA

- Only genes that are expressed in the cells from which the original mRNA is obtained (tissue specific libraries) will be represented
- frequency in the library corresponds to transcription level of the gene
- will only include coding sequences (no introns, no non-genes)
- can be used to express the protein in bacteria

Genomic

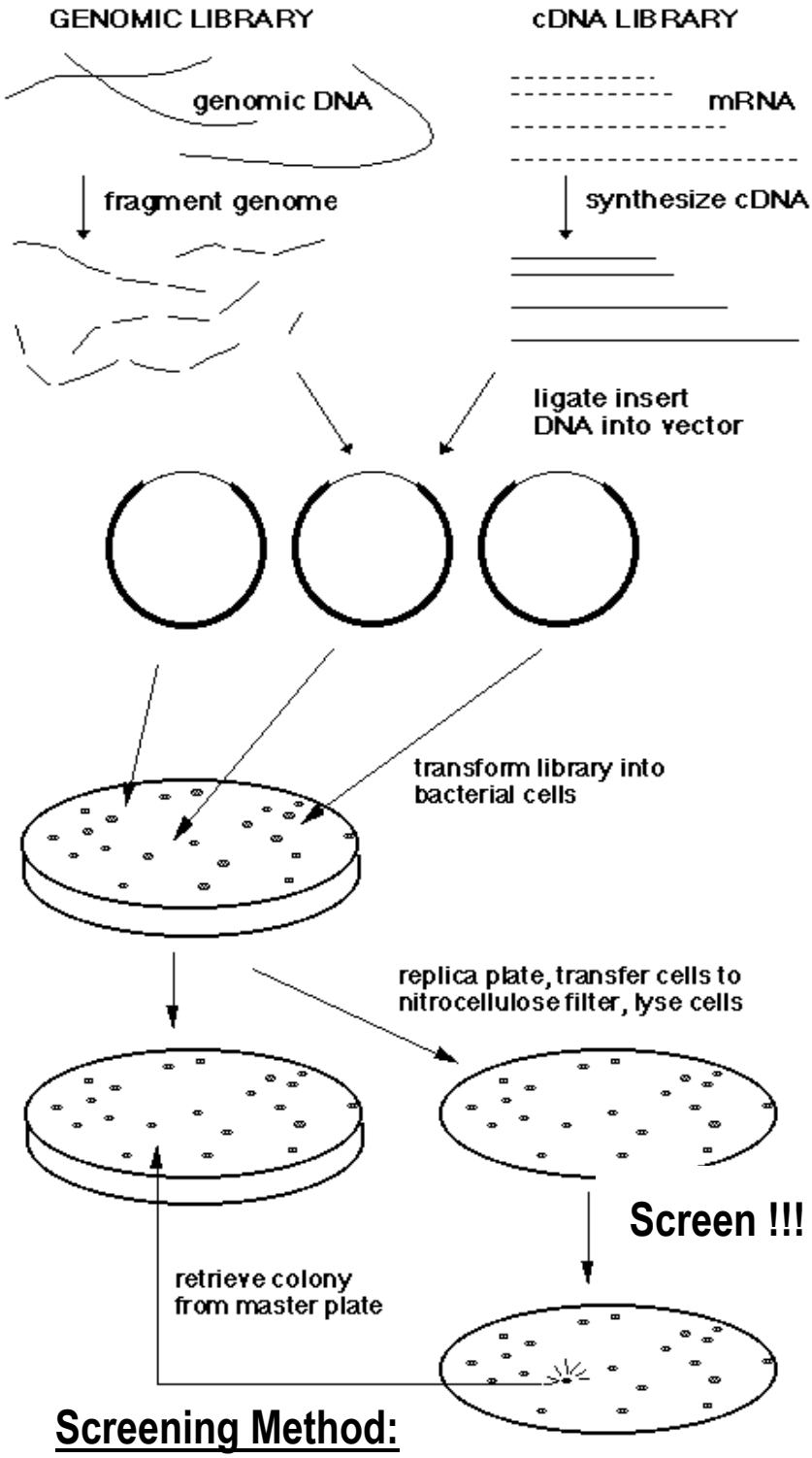
- includes all genes (complete representation of all the sequences)
- includes non-coding sequences (including introns)

- cDNA libraries absolutely necessary if functional eukaryotic protein is needed
- Genomic libraries are used for sequencing projects or when working with bacterial species (cDNA library offers no advantage)
- **Expression Libraries** are made by using a vector that incorporates a promoter upstream of the cloning site; protein expression can be induced
  - Eukaryotic: cDNA
  - Prokaryotic: Genomic

Library Screening Strategies:

- Three major scenarios for screening a library for the sequence (or gene) of interest
  - **A) DNA (or protein) sequence or related sequence is already known**
    - part of the gene is already known
    - gene is already known in another organism
  - **B) An antibody to the protein is available**
  - **C) The protein of interest can be detected or selected for**
    - assay available for protein activity

Library Screening Strategies (cont):



Choice of Library:

A) Genomic or cDNA
B) Expression Library
C) Expression Library

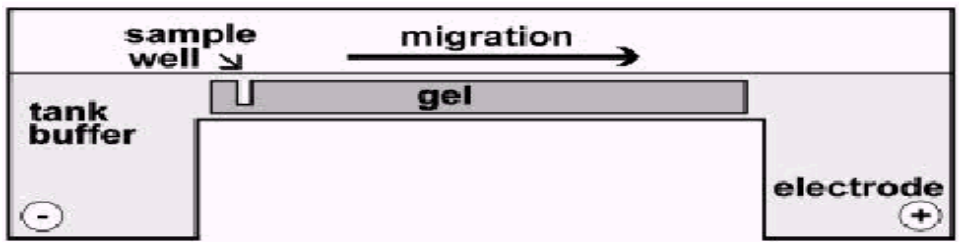
Screening Method:

A) DNA Hybridization to gene-specific probe
B) Detection of protein with antibody
C) Detect protein activity using assay

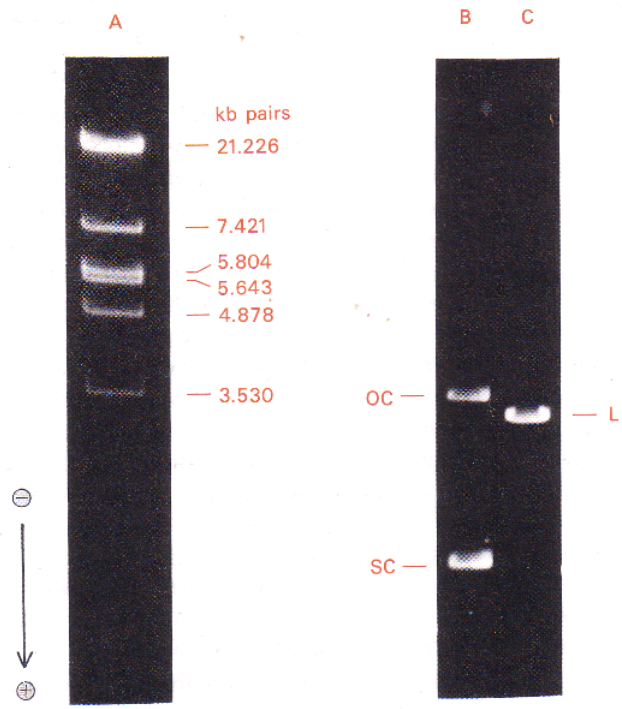
Analytical Tools for DNA, RNA, Protein: electrophoresis & blotting

Gel Electrophoresis

- macromolecules can be separated by using electrical fields
  - DNA and RNA have negative charge
  - Protein can be treated in SDS to have a negative charge
- samples are run in polymeric "gels" (agarose, polyacrylamide)
- constant mass to charge ratio
  - separation according to size & 3-D conformation
  - "snaking" through gel
  - larger molecules run slower

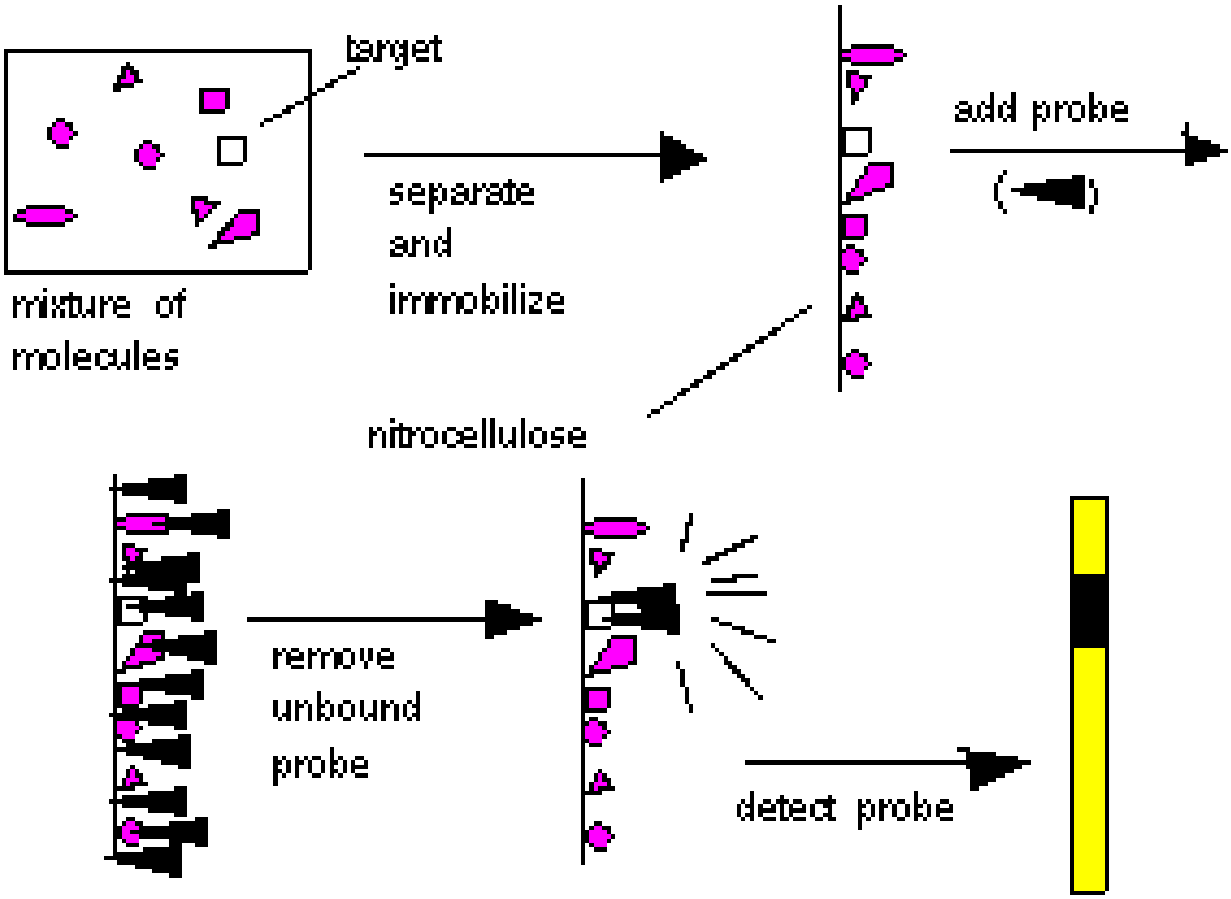


- After separation, specific stains can be used to visualize DNA, RNA, or Protein
  - DNA & RNA: ethidium bromide, Sybr Green…
  - Protein: colloidal coomassie blue, silver stain…



The concept of "Blotting"

- After electrophoretic separation of DNA, RNA, Protein
- Following gel electrophoresis, transfer of macromolecules to solid support
  - nitrocellulose, nylon membranes
- Detection of specific macromolecules ("**Probing**") by complementarity
  - DNA / RNA : using labeled DNA probes
  - Proteins: using antibodies



Blot type	Target	Probe	Applications
Southern	DNA	DNA or RNA	mapping genomic clones estimating gene numbers
Northern	RNA	DNA or RNA	mRNA size, abundance mRNA expression patterns
Western	Protein	Antibodies	protein size, abundance

- Is it there ?
- How much ?
- How big ?

Detection of specific macromolecules by complementarity:

1) DNA-DNA. A single-stranded DNA (ssDNA) probe molecule can form a double-stranded, base-paired hybrid with a ssDNA target if the probe sequence is the reverse complement of the target sequence.

2) DNA-RNA. A single-stranded DNA (ssDNA) probe molecule can form a double-stranded, base-paired hybrid with an RNA (RNA is usually a single-strand) target if the probe sequence is the reverse complement of the target sequence.

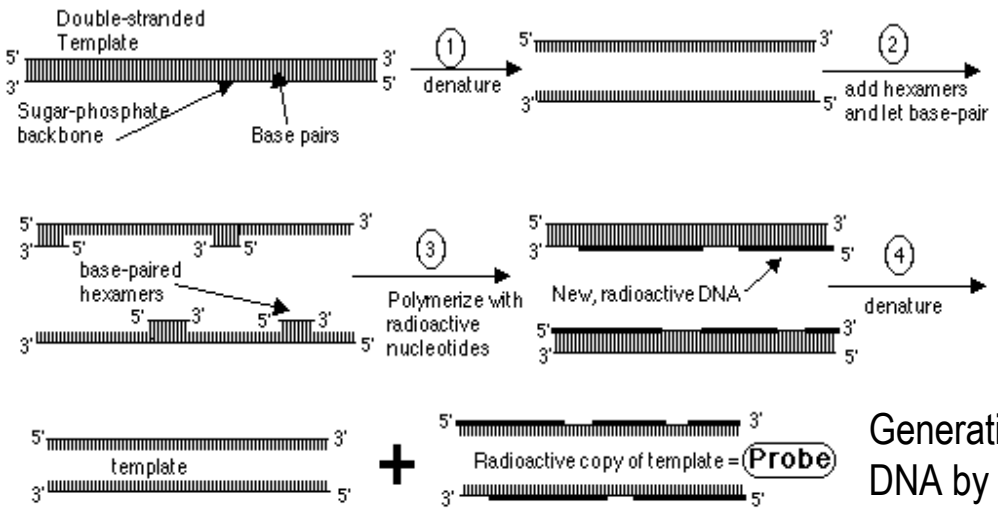
3) Protein-Protein. An antibody probe molecule (antibodies are proteins) can form a complex with a target protein molecule if the antibody's antigen-binding site can bind to an epitope (small antigenic region) on the target protein.

The two most important features of hybridization:

1) Hybridization reactions are highly specific

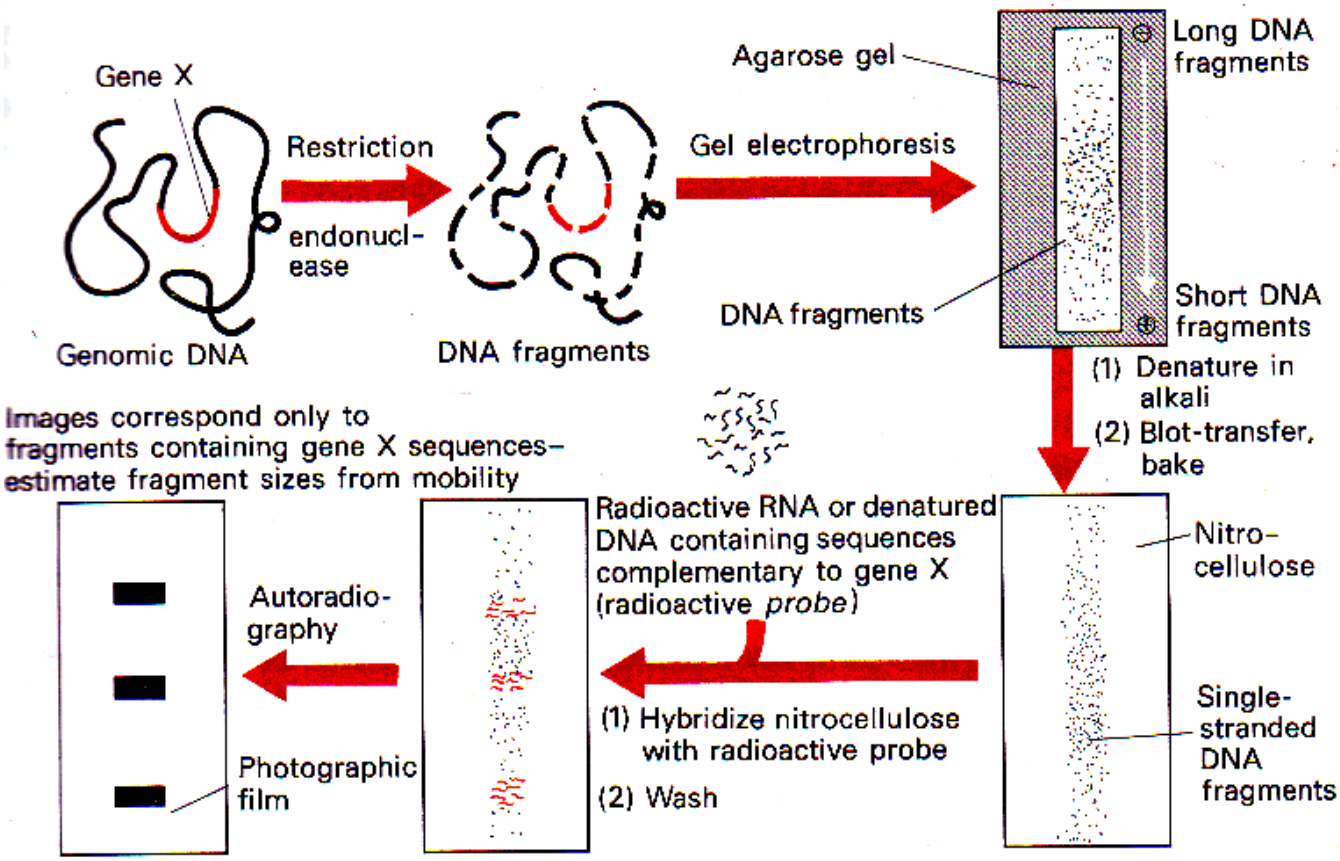
2) Hybridization reactions will occur in the presence of large quantities of molecules that are similar but not identical to the target.

- Cells contain thousands of genes, mRNAs, proteins
- Hybridization can be use to detect specific macromolecules in a complex mixture of other macromolecules

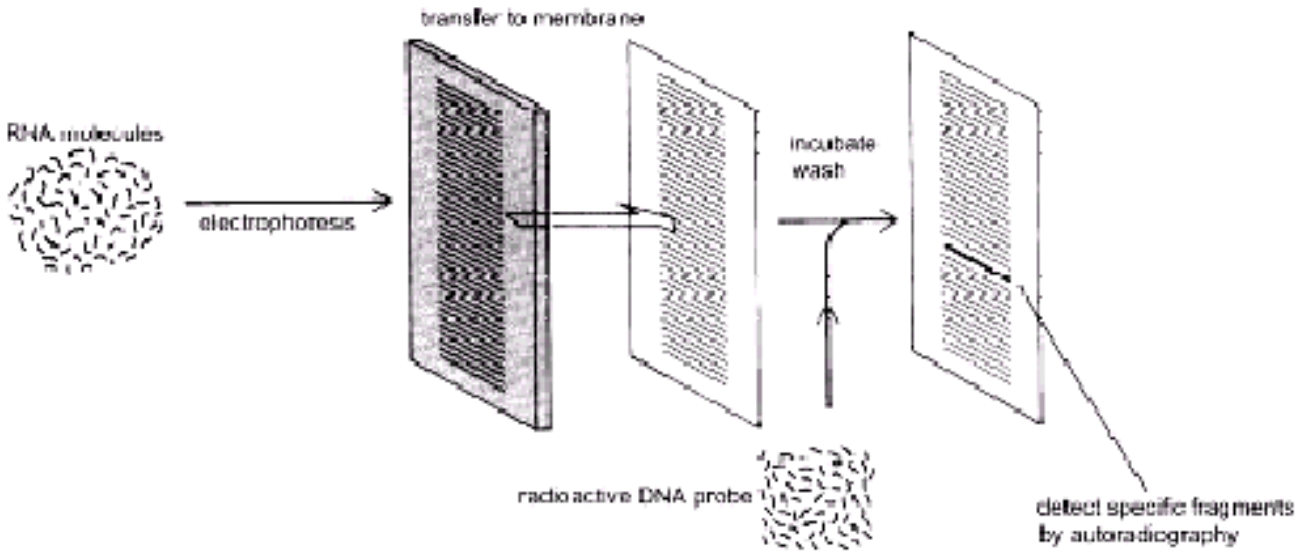


Generating radioactively labeled DNA by random priming

Southern Blotting



Northern Blotting



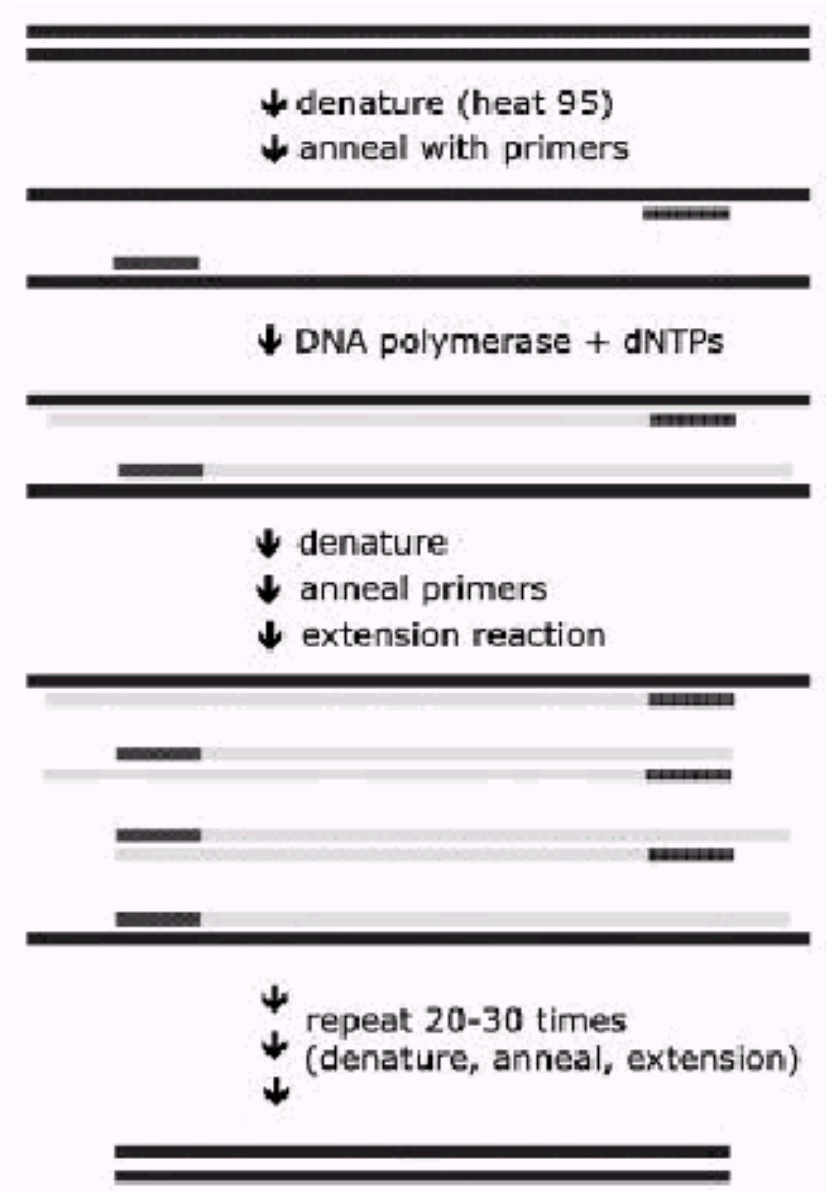
Southerns vs. Northerns vs. Westerns

Southern Blot	Northern Blot	Western Blot
1) Extract DNA from cells	1) Extract RNA from cells	1) Extract protein from cells
2) Cut with restriction enzyme		
	2) Denature with formaldehyde	2) Denature with SDS
3) Run on gel (usually agarose)	3) Run on gel (usually agarose)	3) Run on gel (usually polyacrylamide — called "SDSPAGE")
3.5) Denature DNA with alkali		
4) Transfer to nitrocellulose (usually by capillary action)	4) Transfer to nitrocellulose (usually by capillary action)	4) Transfer to nitrocellulose (usually by electrophoresis)
5) Block with excess DNA	5) Block with excess RNA	5) Block with excess protein
6) Hybridize with labeled DNA probe	6) Hybridize with labeled DNA probe	6) Hybridize with labeled antibody probe
7) Wash off unbound probe (use controlled stringency)	7) Wash off unbound probe (use controlled stringency)	7) Wash off unbound probe
8) Autoradiograph	8) Autoradiograph	8) Autoradiograph or develop with chromogenic substrate

	Southern	Northern	Western
What is separated by molecular weight? (target)	DNA cut with restriction enzymes	RNA denatured with formaldehyde	Protein denatured with SDS
Probe	radioactive gene X DNA	radioactive gene X DNA	Antibody against protein X, labeled with radioactivity or enzyme
What do you learn from it?	Restriction map of gene X in chromosome	-how much gene X mRNA is present? -how long is gene X mRNA?	-how much protein X is present? -how big is protein X?

The Polymerase Chain Reaction

- enzymatic amplification of DNA sequences
- need oligonucleotide primers bracketing the target region of interest
- Thermostable DNA polymerase; dNTPs
- Thermal cycling: denaturing; annealing; extension



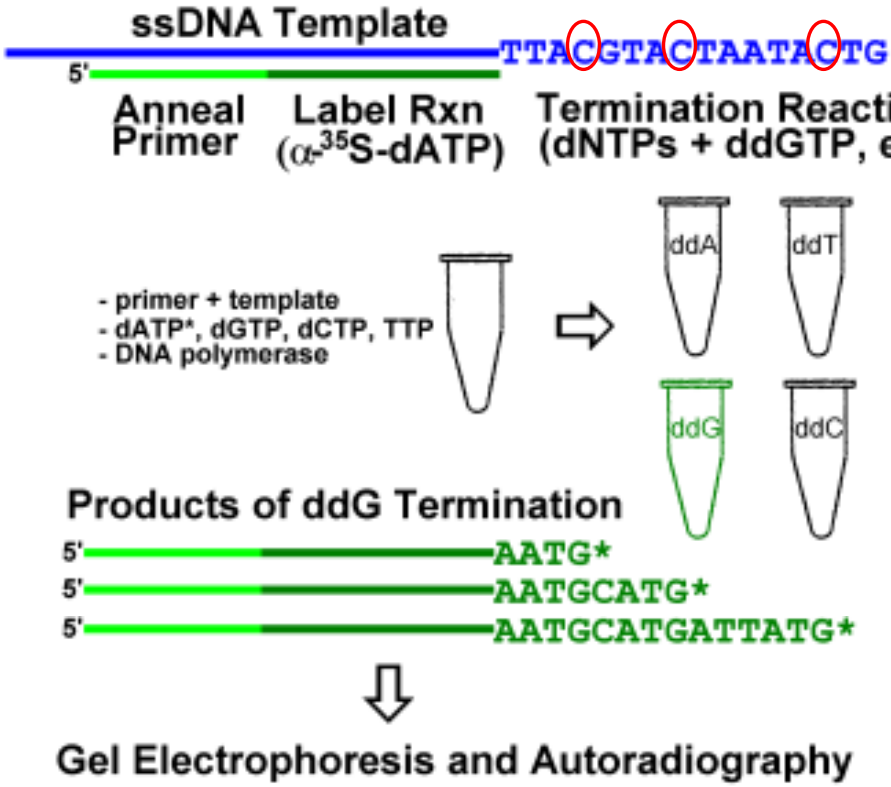
After 30 cycles

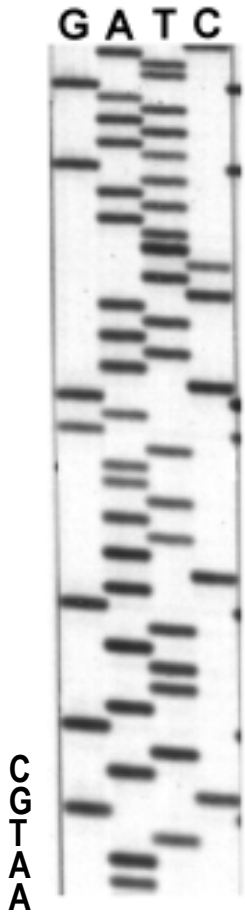
- the amplification is 2^{30}
- $2^{10} \sim 1000$, $2^{30} \sim 10^9$

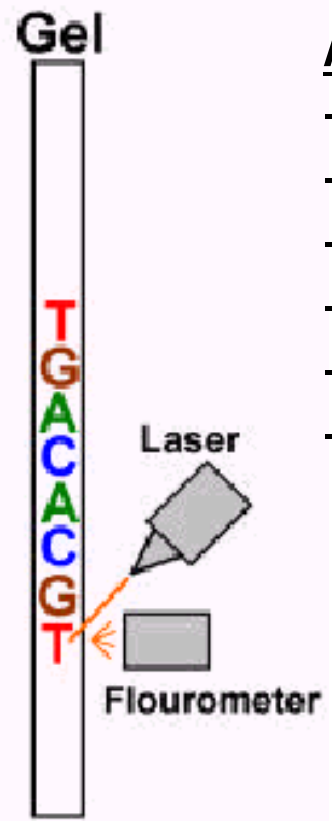
~ billion-fold amplication

- Products from one cycle can become targets in the next cycle
- Exponential amplification of the target region !!!!!
- Amplification without cloning: quick and easy

DNA Sequencing

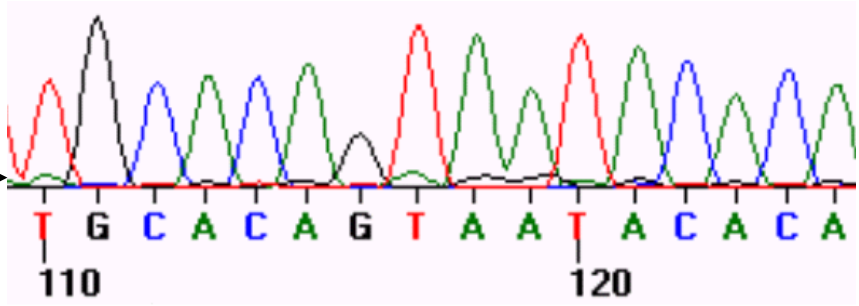






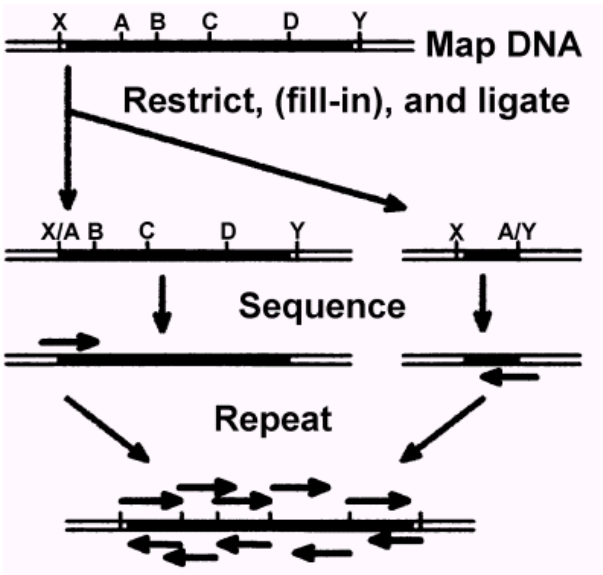
Automated Sequencing:

- same chemistry
- fluorescent dye terminators
- all 4 termination reactions in one tube
- fragments of increasing size pass by detector
- chain terminator detected in real time
- sequence decoded in real time

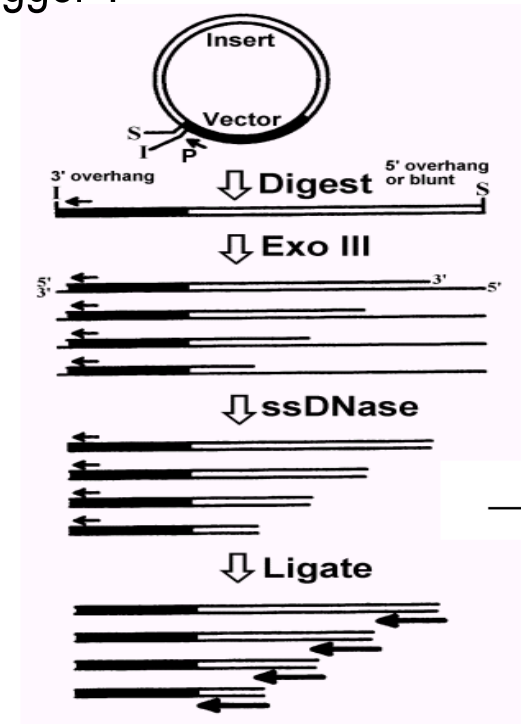


DNA Sequencing Strategies

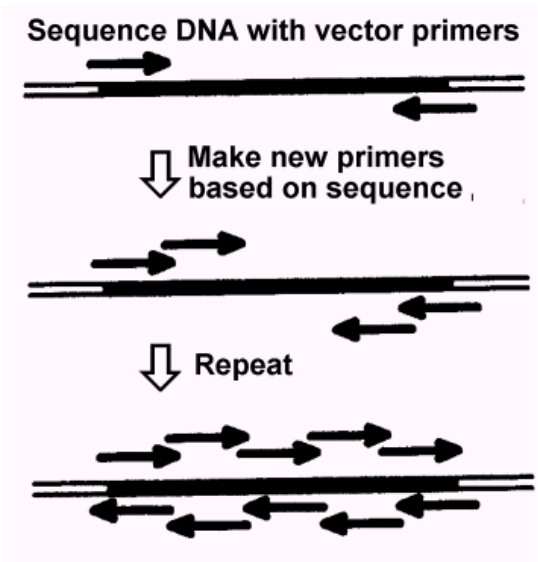
- From one primer / template combo: 500 - 700 bp of sequence
- What to do if you have something bigger ?



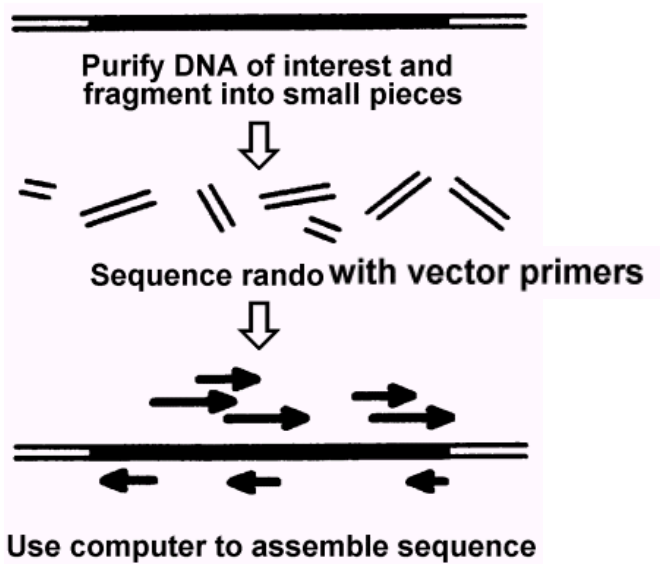
Subcloning



Nested Deletions



Primer Walking



Shotgun Sequencing