

PLNT2530
2025

Unit 6b
Sequence Libraries

Molecular Biotechnology (Ch 4)
Analysis of Genes and Genomes (Ch 5)

To find or isolate a gene or promoter you need a library

Library is a collection of sequences

Libraries are of two general types

Genomic library – goal is to have all the sequence information in the genome represented in the library

cDNA library - a very selective library which is designed so that it will have only the coding sequences of expressed genes represented.

Genomic library

- Organism specific and represents the whole genetic material
- Not tissue and development-specific
- Contain expressed genes, non-expressed genes, exons and introns, promoter and terminator regions, but **mostly repetitive DNA**

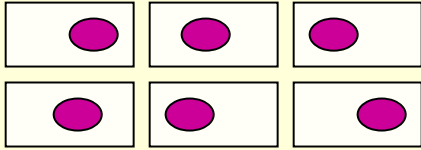
cDNA library

- Sequences are obtained by the conversion of mRNA to cDNA from a specific tissue
- Tissue and developmental-specific
- Varies in abundance (highly expressed genes-multiple times in the library and low expressed genes will be represented less)
- Contains coding sequences including 5' and 3' untranslated regions
- no intergenic regions

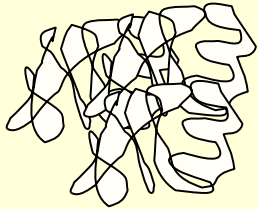
A Genomic Library

- Prepared by cutting the genomic DNA into fragments and select for the appropriate size (based on vector used)
- Insert fragments into an appropriate vector, introduce into bacteria, segregate and amplify each fragment.

Genomic Library



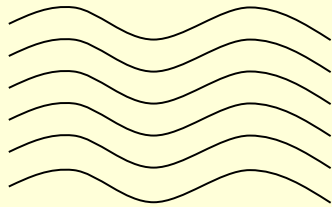
Tissue composed of 1000's of cells containing nuclei with complete sets of chromosomes



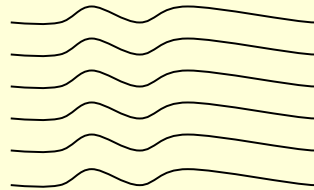
Break cells and isolate DNA

1000's of copies of genomic DNA

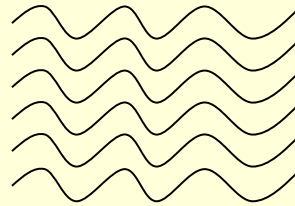
Chromosome 1



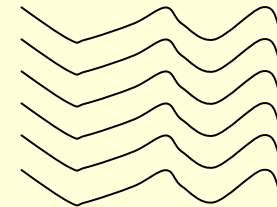
Chromosome 2



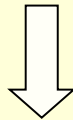
Chromosome 3



Chromosome n



1000's of copies of genomic DNA

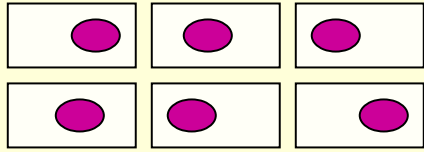


Partial digestion with restriction enzyme

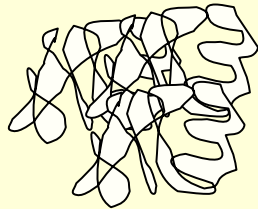


Required size range

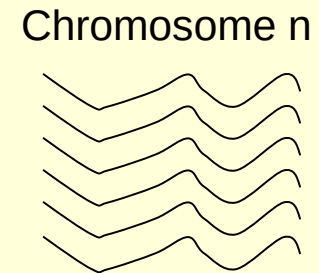
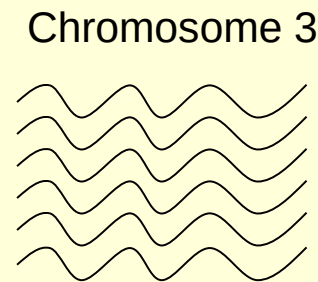
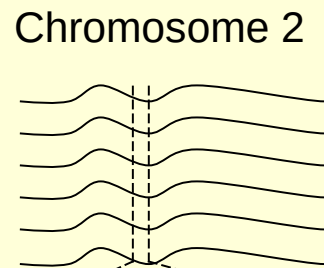
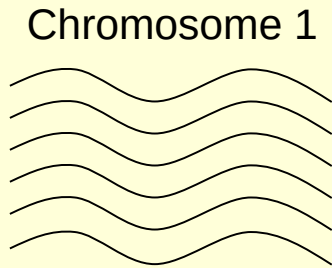
Genomic Library



Tissue composed of 1000's of cells containing nuclei with complete sets of chromosomes

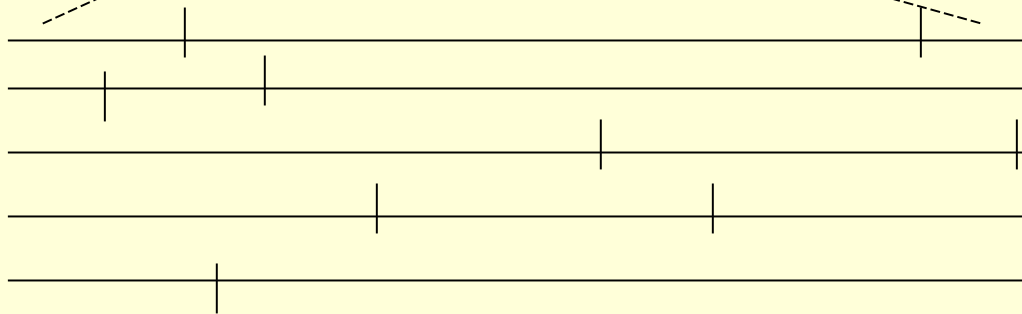


Break cells and isolate DNA
1000's of copies of genomic DNA



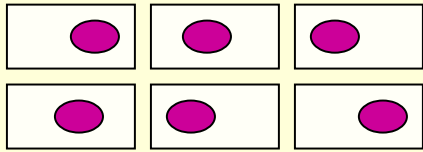
1000's of copies of genomic DNA

Short digestion

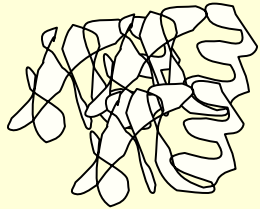


Required size range

Genomic Library



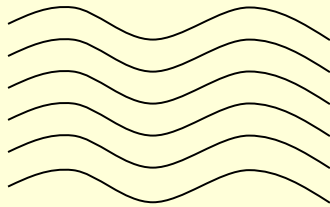
Tissue composed of 1000's of cells containing nuclei with complete sets of chromosomes



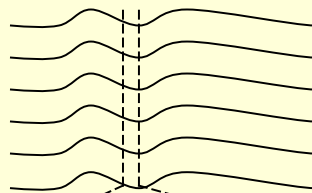
Break cells and isolate DNA

1000's of copies of genomic DNA

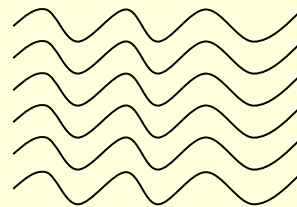
Chromosome 1



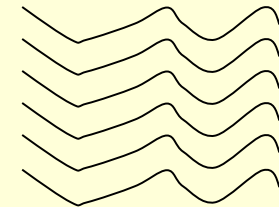
Chromosome 2



Chromosome 3

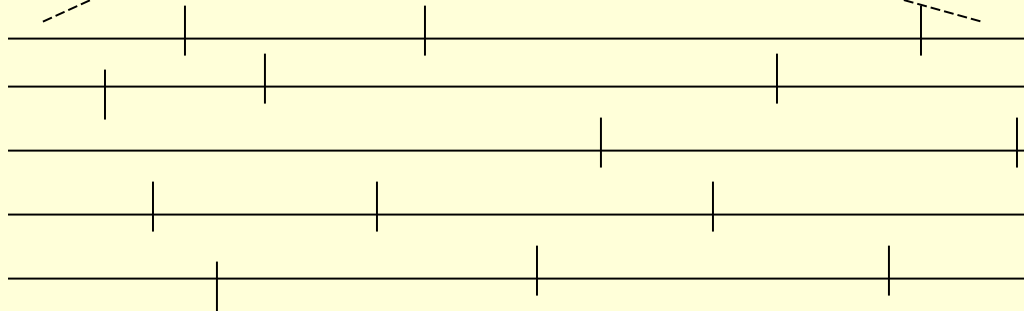


Chromosome n



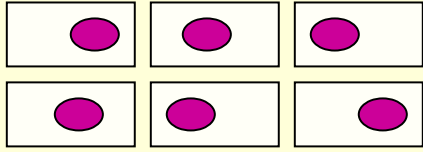
1000's of copies of genomic DNA

longer digestion

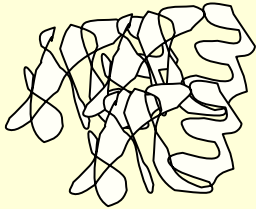


Required size range

Genomic Library



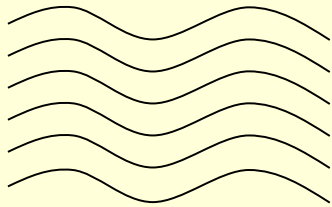
Tissue composed of 1000's of cells containing nuclei with complete sets of chromosomes



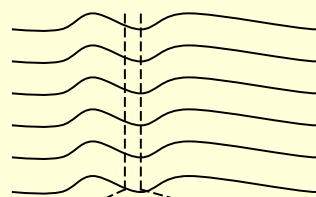
Break cells and isolate DNA

1000's of copies of genomic DNA

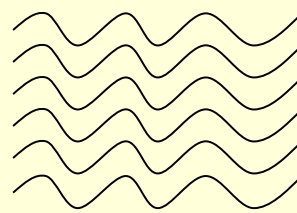
Chromosome 1



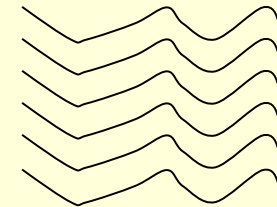
Chromosome 2



Chromosome 3

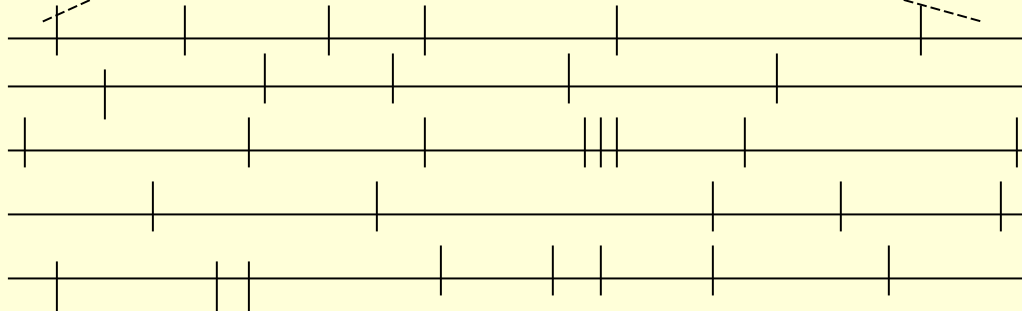


Chromosome n



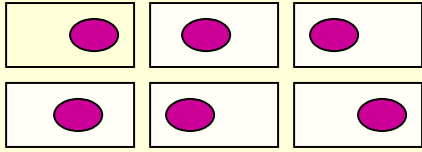
1000's of copies of genomic DNA

longer digestion

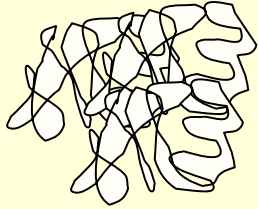


Required size range

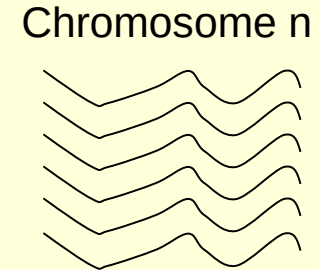
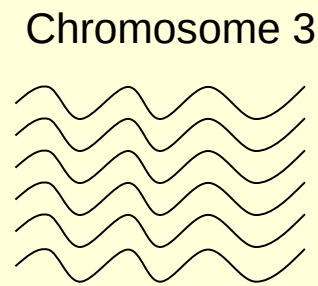
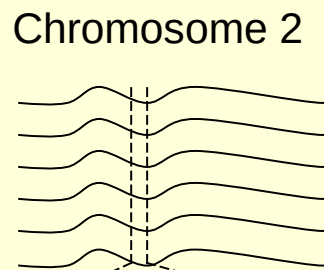
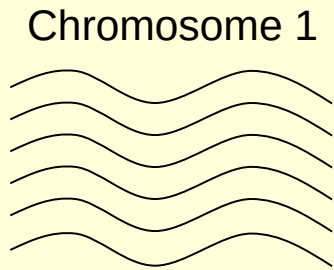
Genomic Library



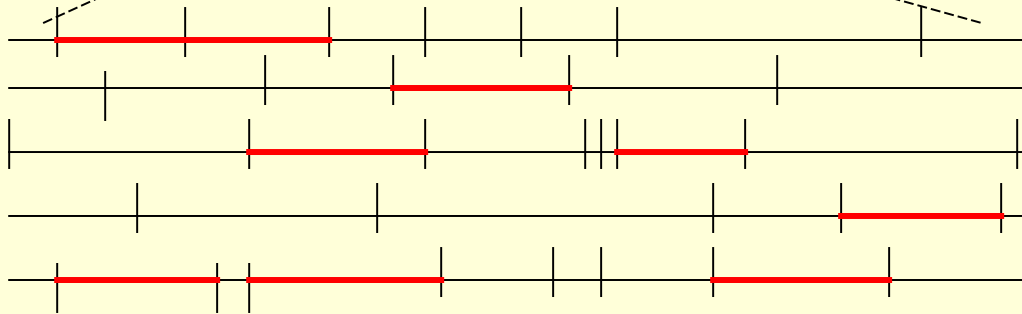
Tissue composed of 1000's of cells containing nuclei with complete sets of chromosomes




Break cells and isolate DNA
1000's of copies of genomic DNA

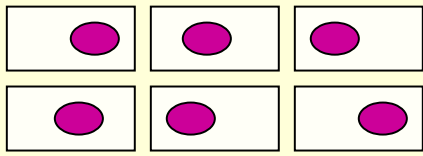


1000's of copies of genomic DNA

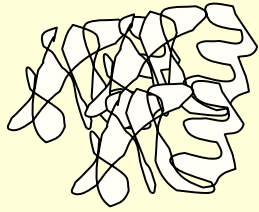



Required size range

Genomic Library

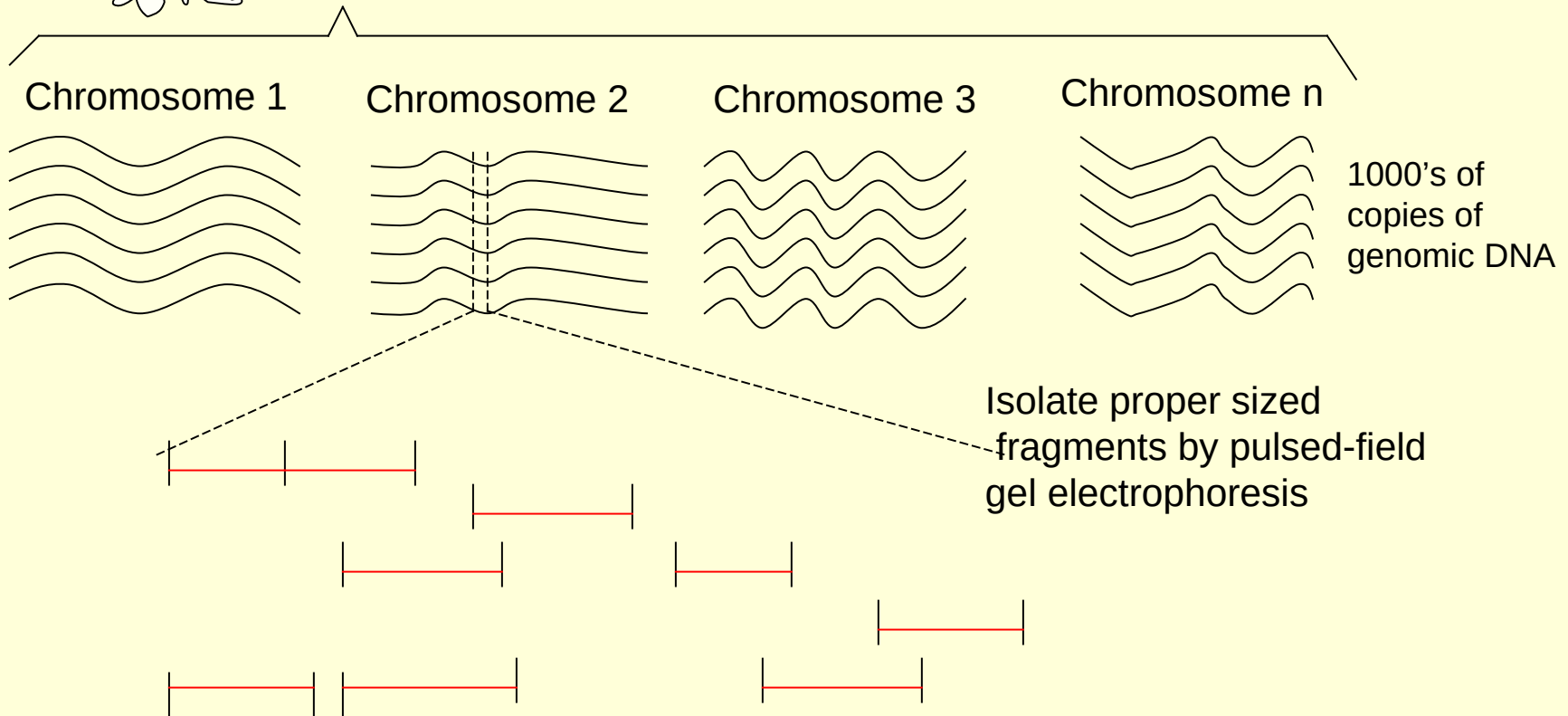


Tissue composed of 1000's of cells containing nuclei with complete sets of chromosomes



Break cells and isolate DNA

1000's of copies of genomic DNA



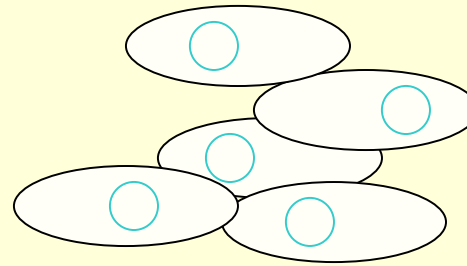
Ligate all pieces into BAC vector, transform bacteria for segregation and amplification

Bacterial transformation

Recombinant BAC vectors
with different genomic DNA
inserts (~100,000 – 250,000 bp)

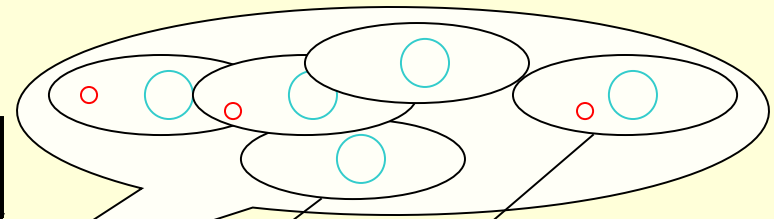
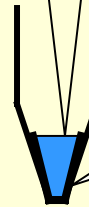


+

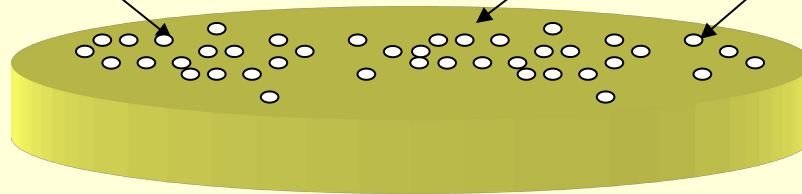


Bacterial cells sensitive
to antibiotic

electroporation (electric pulse temporarily
makes membranes porous)



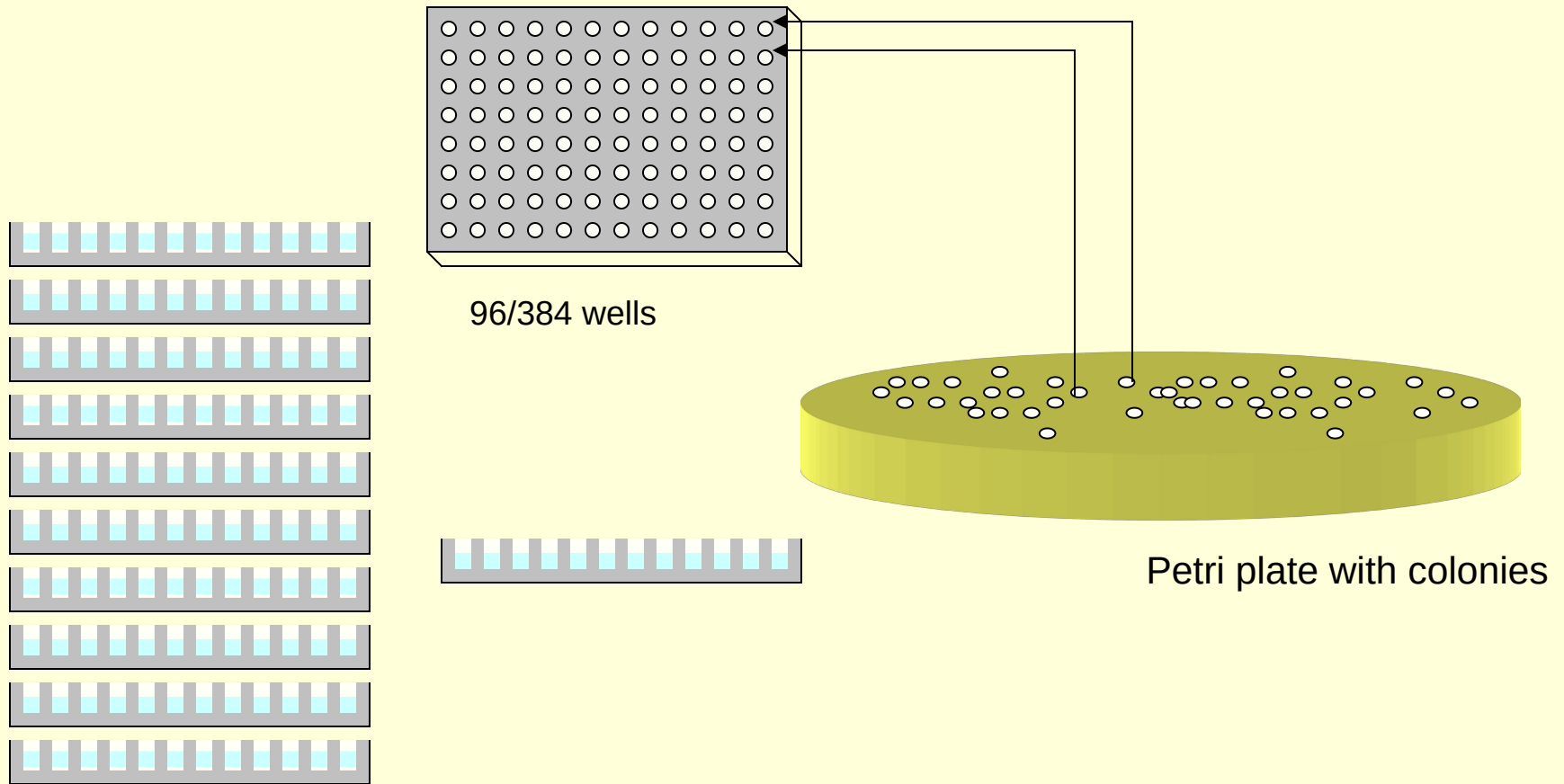
Bacterial colony derived
from a single cell



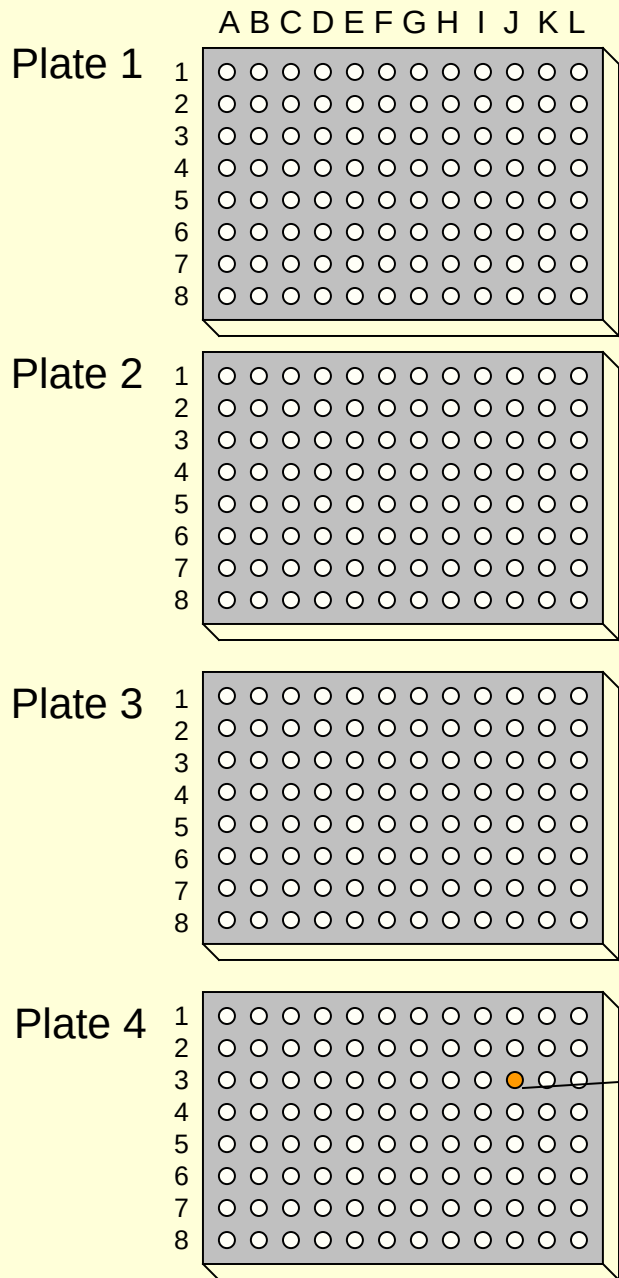
Bacterial colonies each
with a BAC vector carrying
a piece of genomic DNA

Genomic library storage

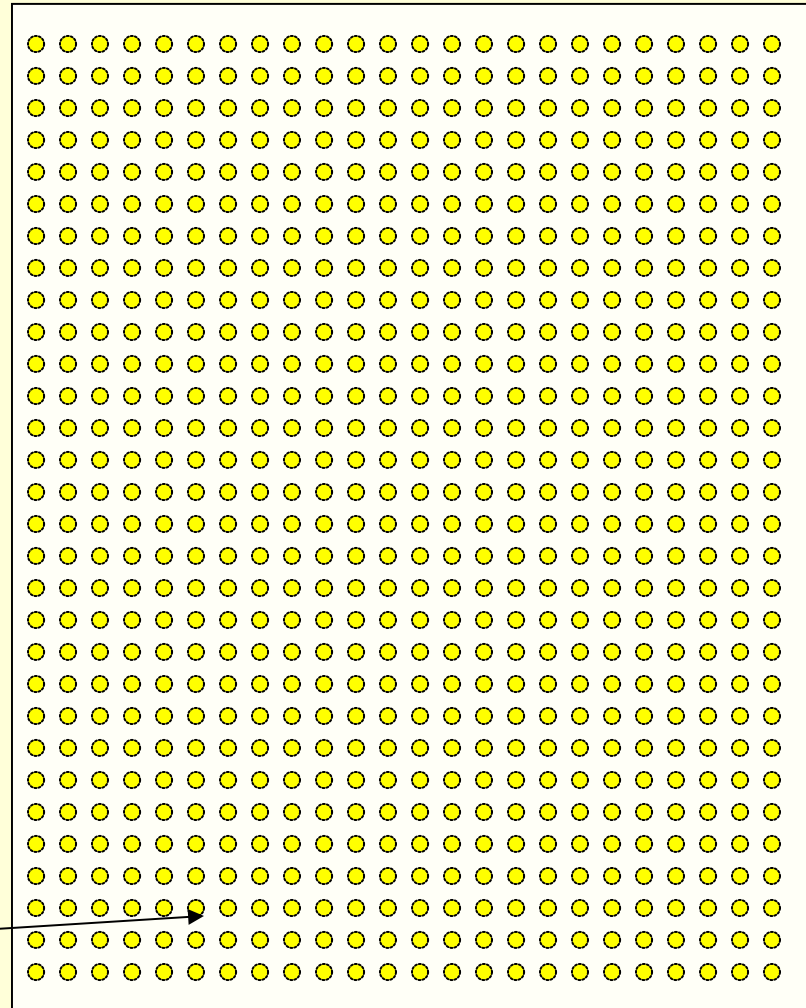
A few cells transferred into growth media
-allowed to grow for 20 h, then stored frozen



Collection of all colonies from many plates forms a library of sequences. Each colony possesses a BAC with a separate piece of genomic DNA

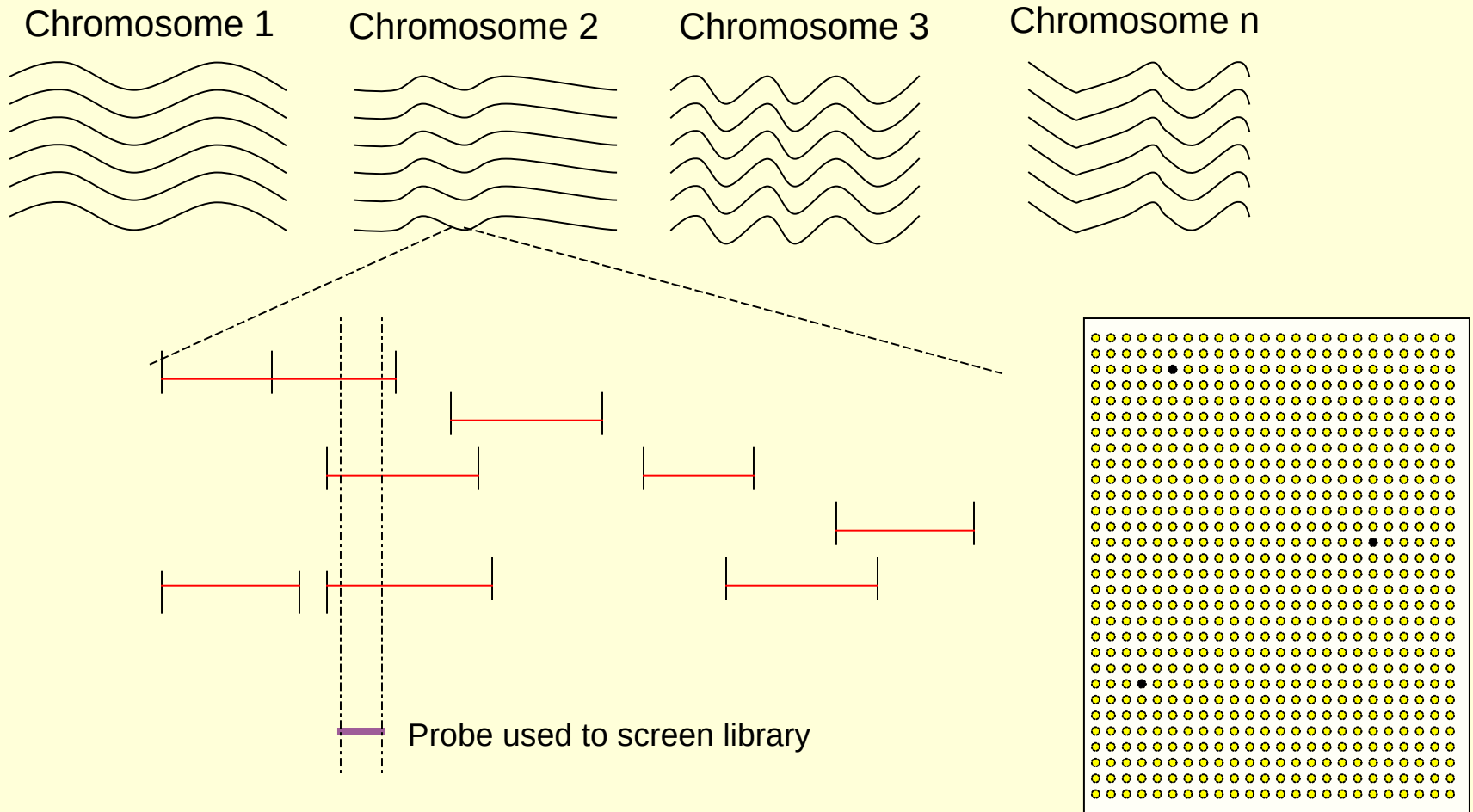


Spot DNA from each sample onto a membrane, denature it, and **screen** for gene of interest



Each spot corresponds to a unique well on one plate

Multiple spots normally appear from library screening

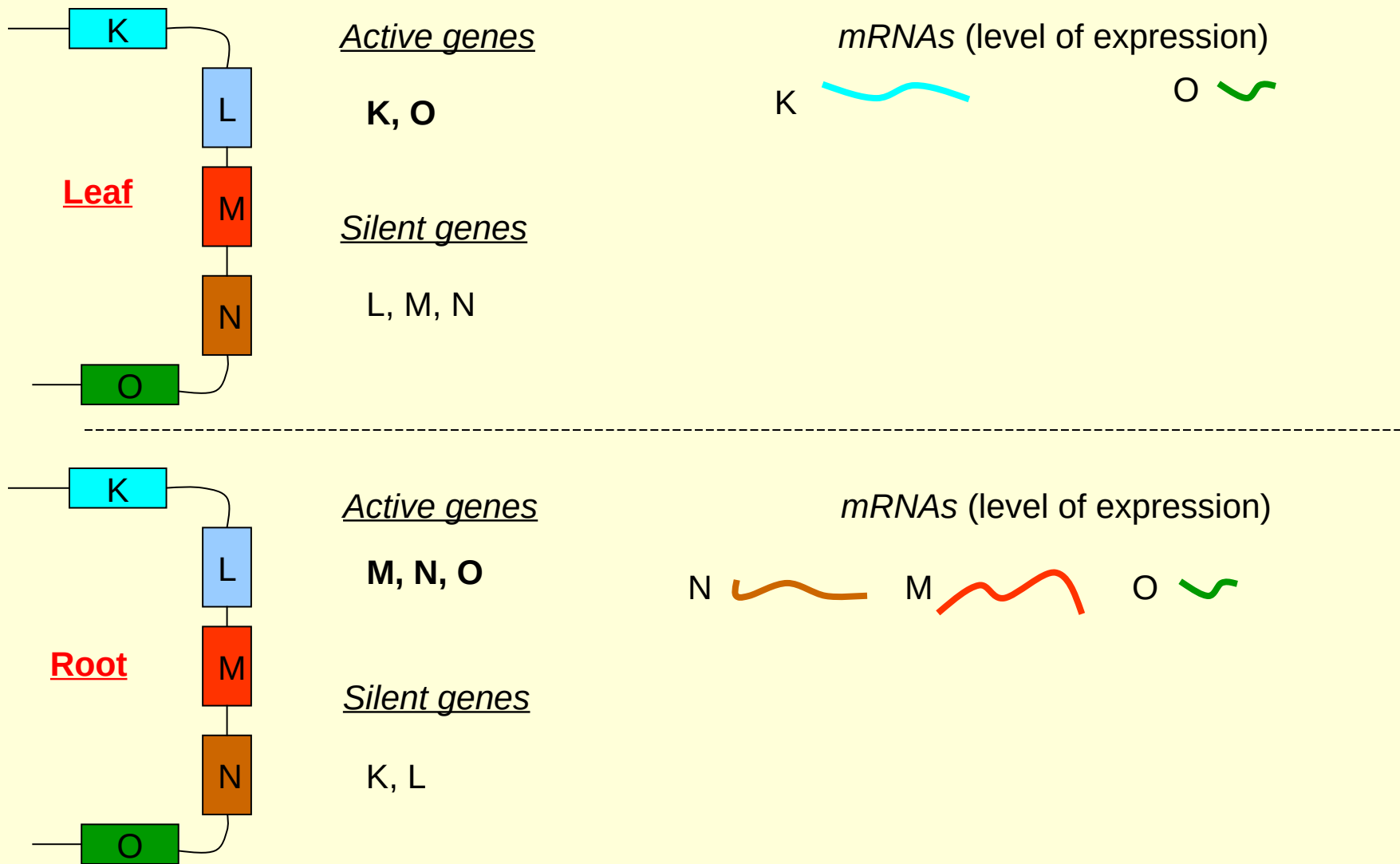


Multiple spots may arise - from redundancy in library due to overlapping fragments or
- from redundancy in the genome of the sequence

cDNA libraries

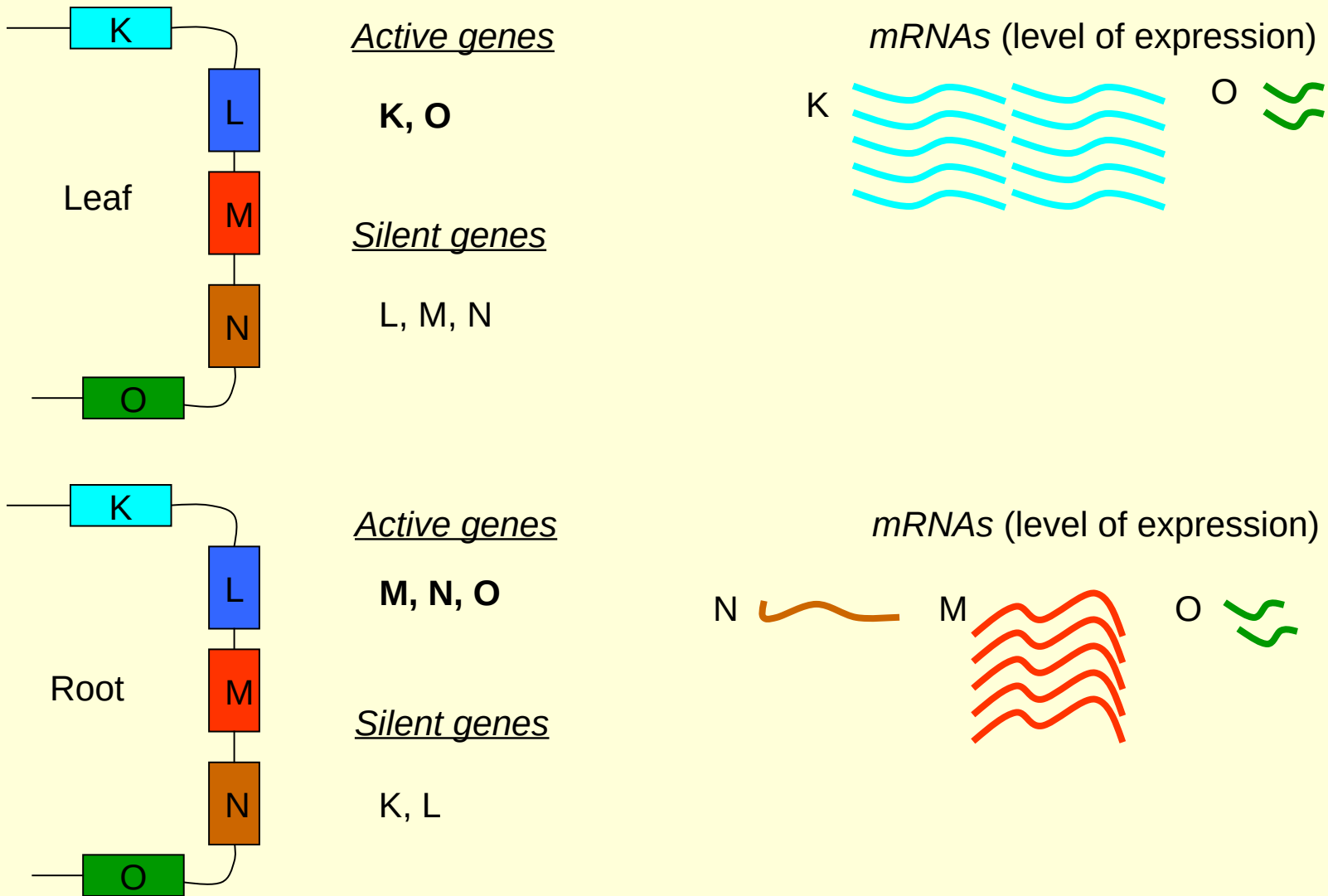
- Made from mRNA by reverse transcription of mRNA into **complementary DNA (cDNA)**
- Contain only processed transcribed sequences
(no regulatory elements, no introns)
- Since transcribed sequences represents only small percentage of genome, only the small percentage of the plant genome that is transcribed will appear in the cDNA library
- Not all genes are expressed in any one tissue
Hence only a fraction of total gene sequences will be present in any cDNA library made from a few specific tissues
- Thus a cDNA Library is far more selective if we are looking for a specific gene than is a Genomic Library

Different cells express different sets of genes



Expressed genes will be expressed at different levels

Different cells express different sets of genes



Different cells express different sets of genes

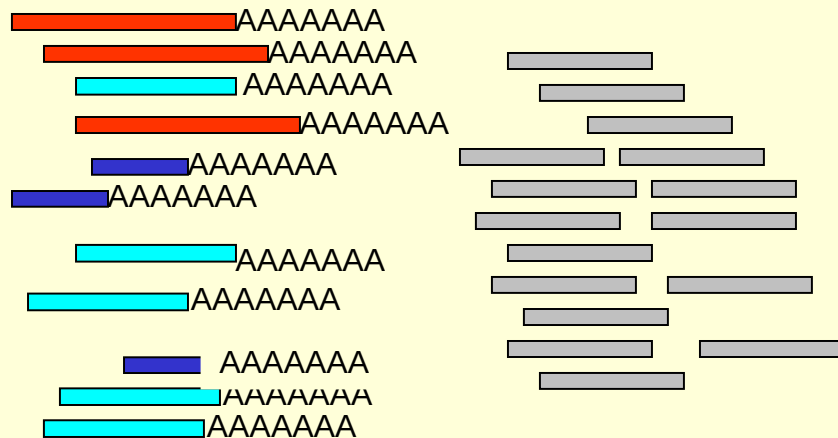
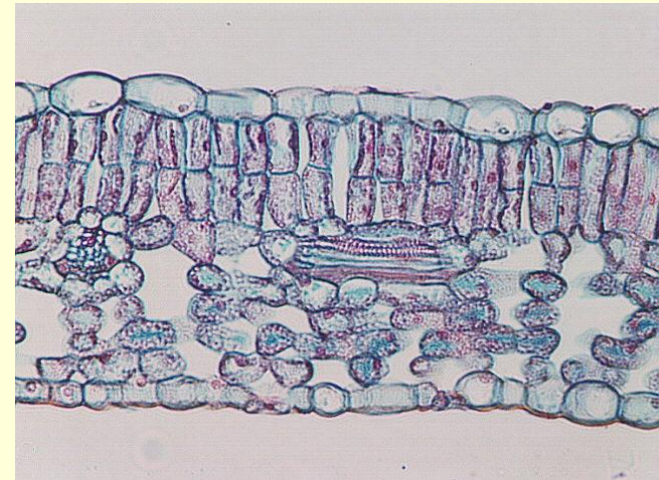
- All cells express some of the same genes (housekeeping genes, central metabolism)
- Different cell types will express different subsets of genes
- At different stages of development different genes are expressed
- Cells respond to specific changes by turning certain genes ON and other genes OFF and by modulating the expression of ON genes

cDNA library construction

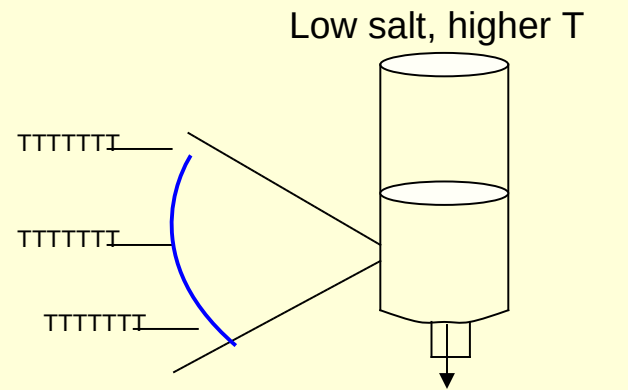
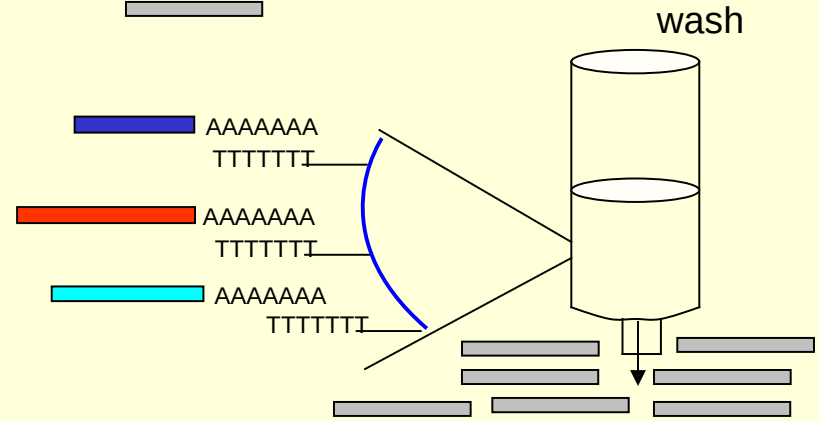
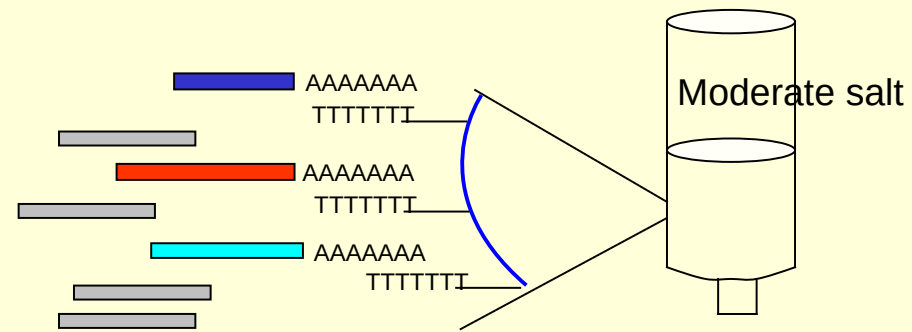
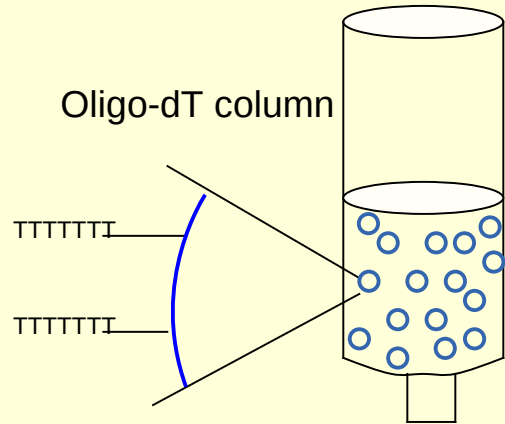
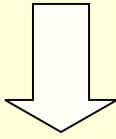
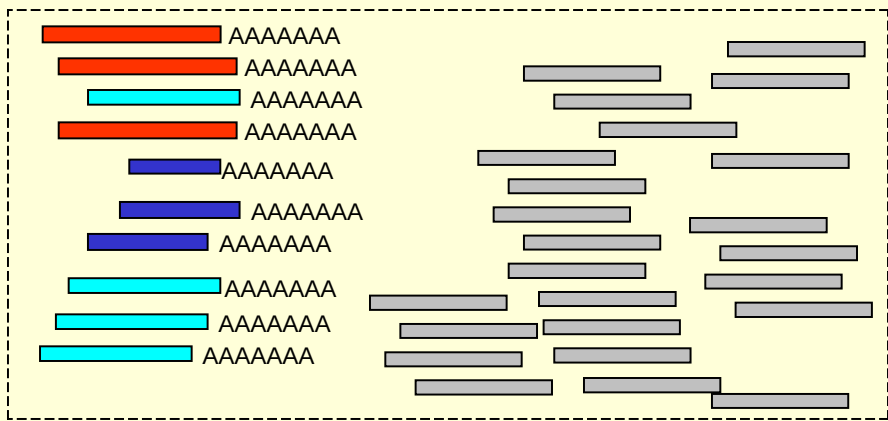
- normally for isolation of specific gene
- mRNA is prepared from selected tissue, at a defined time
 - Based on knowledge of **where** and **when** your gene is expressed

Note: Because most tissues have many different cell types, the mRNA pool will include mRNAs from a diversity of cell types

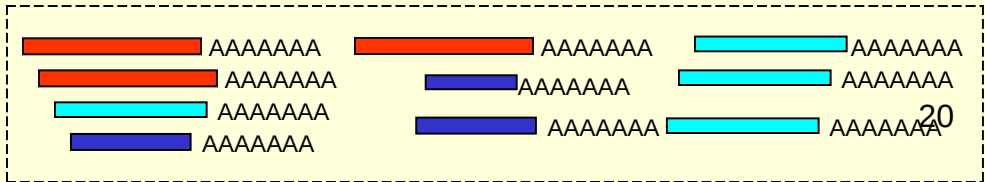
Break cells open and extract RNA



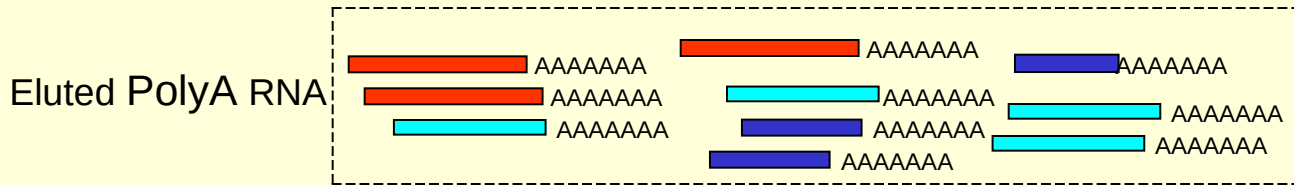
Mixture of rRNA (95%) and mRNA (1-5%)
mRNAs will be a **population** of different gene transcripts that were present in the cells when they were broken



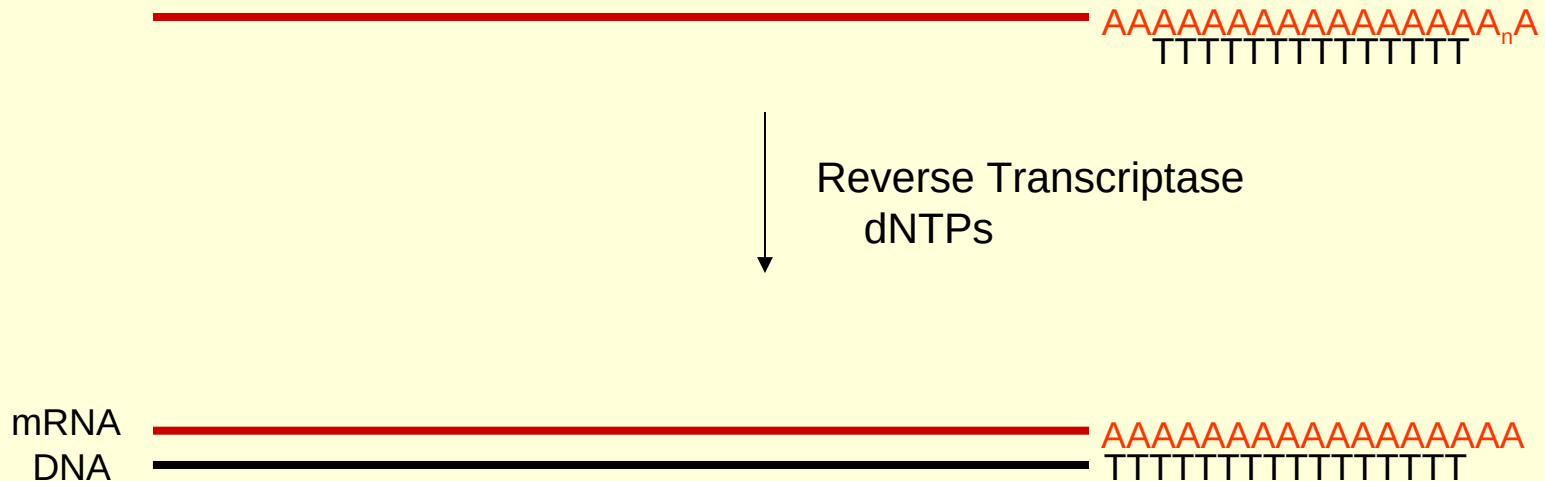
Eluted PolyA RNA



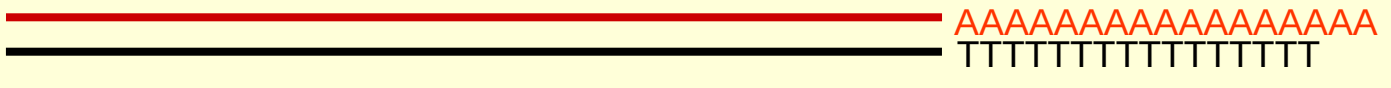
Preparation of cDNA from mRNA



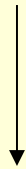
Using Reverse Transcriptase (RNA template-dependent DNA polymerase) and oligo-dT (15-18mer) primer, mRNA copied into complementary DNA (**cDNA**)



mRNA
DNA



AAAAAAAAAAAAAAAAAAAA
TTTTTTTTTTTTTTTTTT

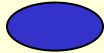


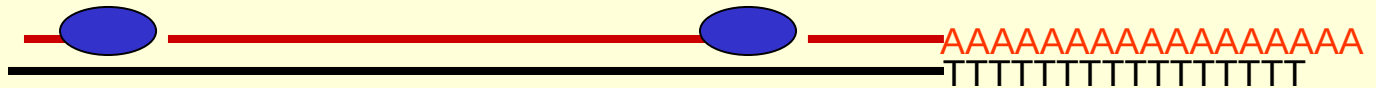
Digest with low level of RNase H to create nicks
in RNA strand to form start points for DNA polymerase



AAAAAAAAAAAAAAAAAAAA
TTTTTTTTTTTTTTTTTT



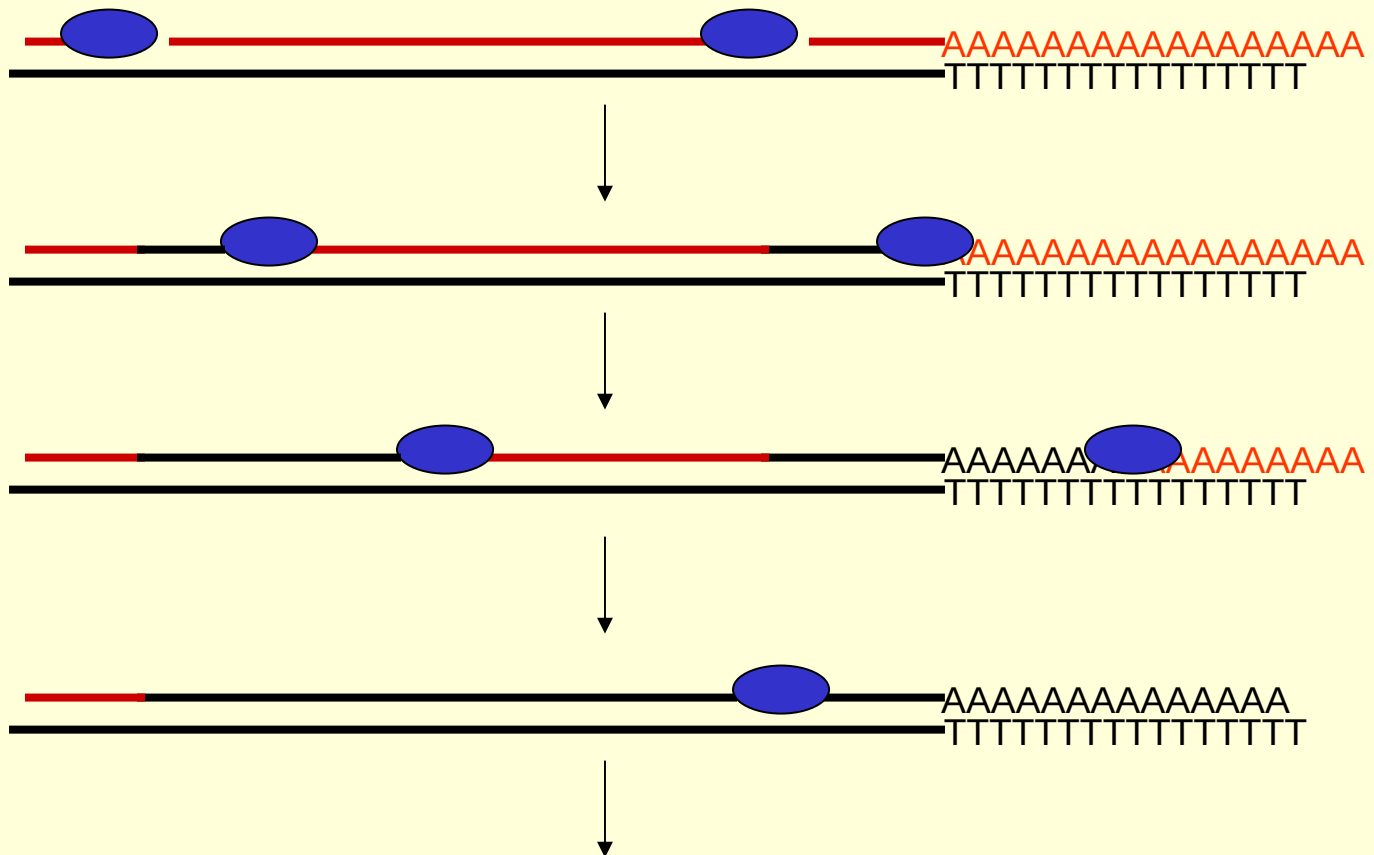
DNA polymerase I  + dNTPs

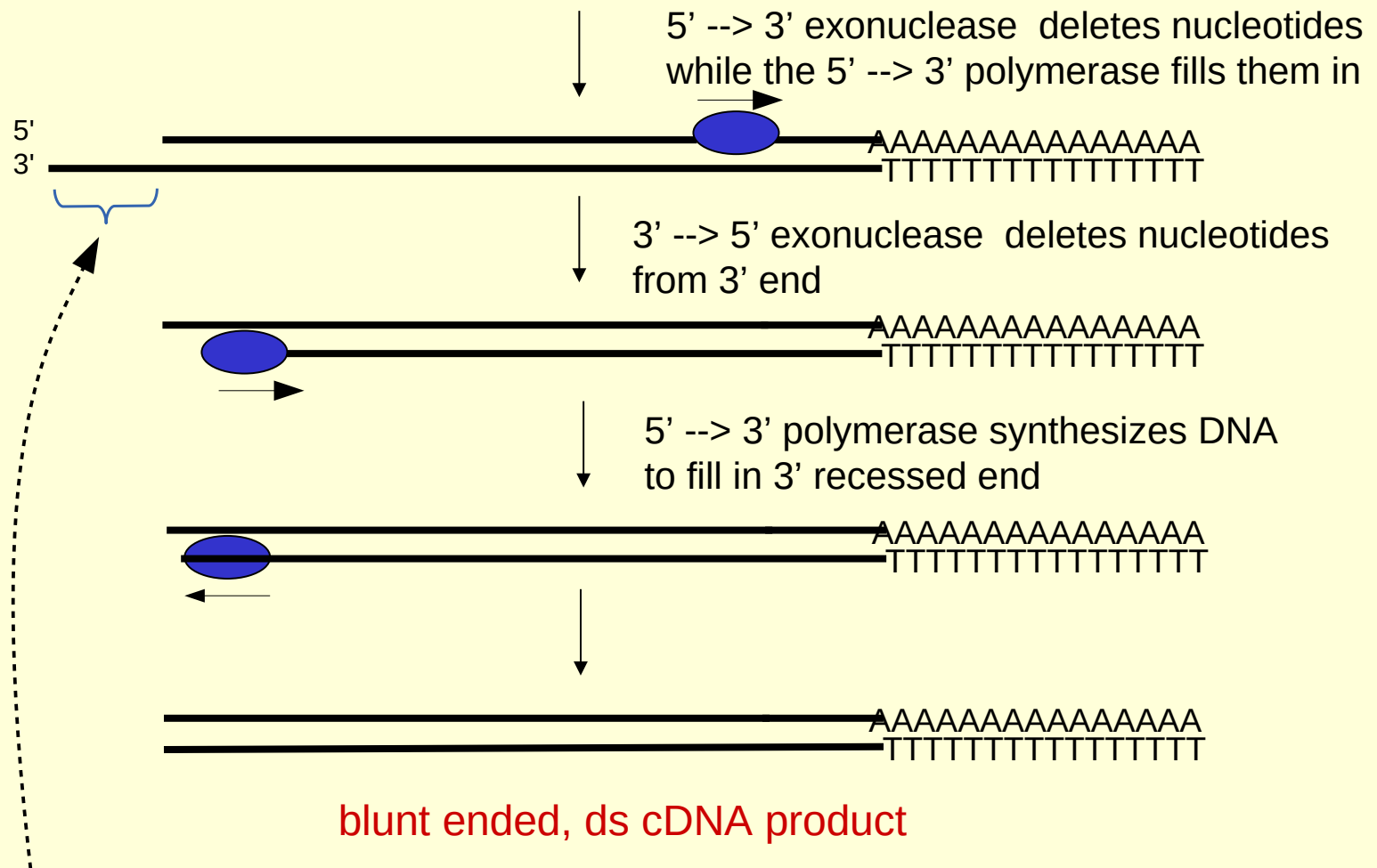


AAAAAAAAAAAAAAAAAAAA
TTTTTTTTTTTTTTTTTT

DNA polymerase I removes ribonucleotides from the RNA strand using the 5' to 3' exonuclease. The 5' to 3' polymerase follows behind, replacing ribonucleotides with deoxyribonucleotides, in a process called "nick translation".

DNA polymerase I + dNTPs





Note: the sequence at the 5' end of the mRNA strand is never replicated, because there is no primer/template complex upstream to copy it. That leaves a 3' protruding end, which is deleted by DNA polymerase. Thus all cDNA clones are missing some of the 5' UTR, and often, some of the coding region of the gene.

Adapters

- A pair of DNA oligomers of different length which are complementary to one another in such a way that they can base pair to create a blunt end at one end of the duplex and a defined restriction enzyme recognition site at the other end. 5' cohesive end is non-phosphorylated while the terminal 5' nucleotide at blunt end is.

5'^PCTAGATTGCG^{3'}

12mer

CTAGATTGCG

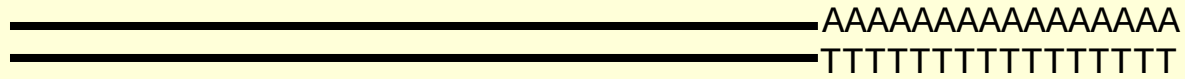
GATCTAACGCTTAA

^{3'}GATCTAACGCTTAA^{5OH'}

16mer

- An adapter pair can be created for any restriction enzyme site.
- Use to create cohesive ends on DNA to improve ligation efficiency

Total population of cDNA molecules



Trim to create blunt ends
Blunt end ligation is **not** efficient
at low concentration

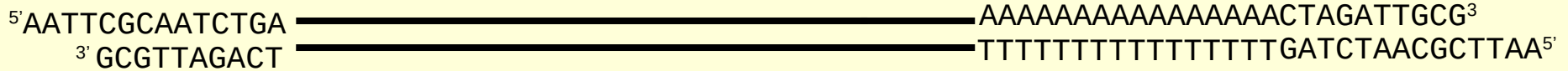


eg. *EcoRI*-adapter pair

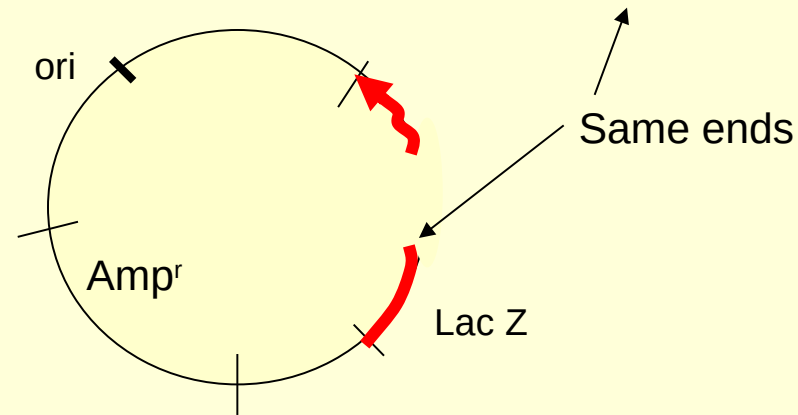
Add excess **adapters** and ligate



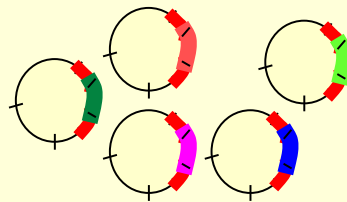
Phosphorylate ends



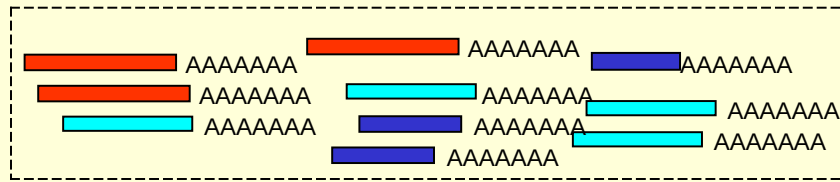
Add dephosphorylated, *EcoRI* linearized
vector and ligate



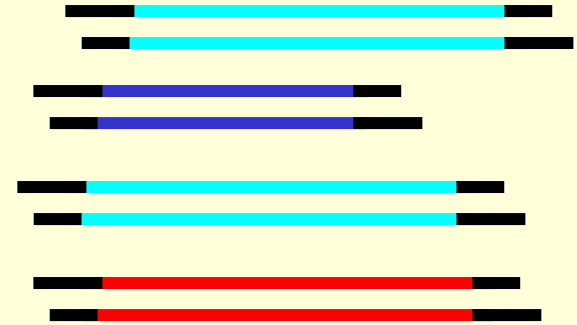
Population of recombinant
vectors with different inserts.



cDNA Library

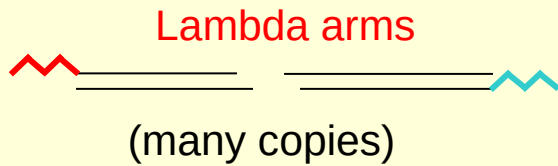


mRNA population

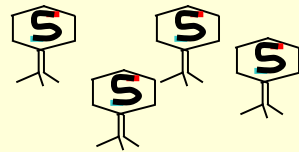


cDNA population

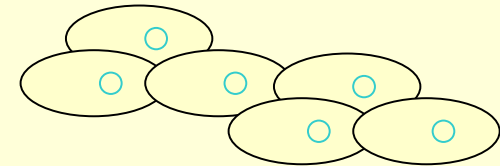
Ligation



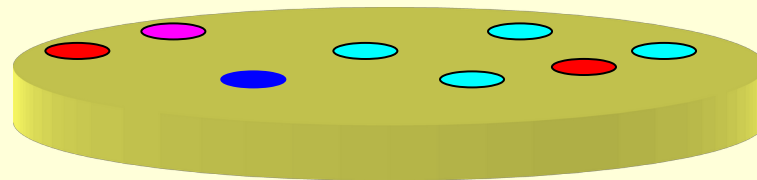
In-vitro packaging of DNA with phage proteins



Infection



Plating



Each plaque has a different cDNA insert
Population of plaques with inserts is a library of gene sequences

- cDNA library can be prepared in either plasmids or phage vectors.
- Normally a cDNA library is screened soon after it is created. cDNA libraries tend to have short shelf lives
- A plasmid based library is preserved by first amplifying the unsegregated library then mixing the bacterial culture with glycerol, freezing rapidly and storing at -80°C
- Phage based libraries are amplified by growing a phage-infected bacterial culture and harvesting the phage particles after cell lysis. The phage can be stored with a drop of chloroform for months at 4°C .