Lecture 3

Tools for studying DNA

blots

PCR

cloning

Northern hybridization (RNA blot) analysis

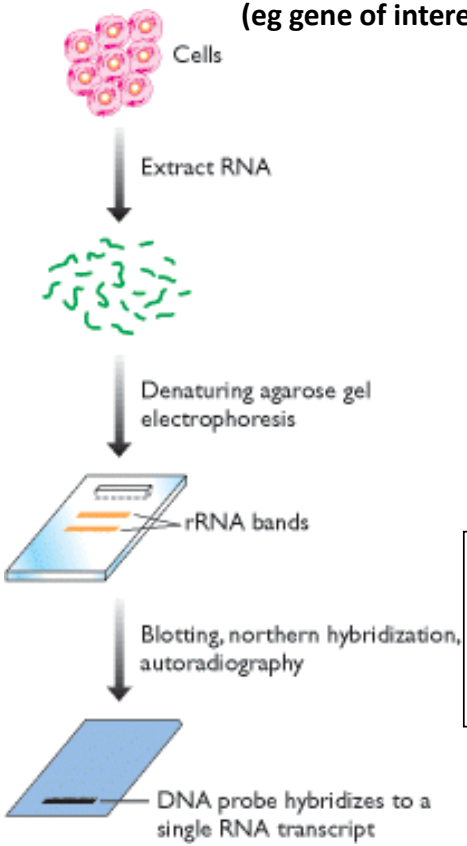
Developed in 1977 @ Stanford

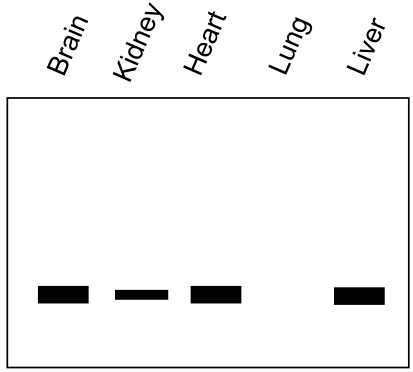
RNA version of Southern

- to identify expressed regions of genomes

– **RNA is electrophoresed, blotted to membrane and hybridized with probe (eg gene of interest)**

- are transcripts are present or not?
- length of mRNA?
- abundance of mRNA?

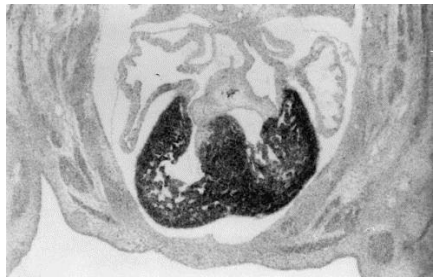




gene X probe

Probe: tagged *DNA (eg. PCR product, restriction fragment, cDNA clone…) in denatured form or oligomer or antisense (synthetic) RNA …*

In situ hybridization
- to determine cellular Location of mRNA



^{35}S-labeled β-myosin antisense probe hybridizing to heart ventricle in 13-day embryonic mouse

Strachan & Read Fig. 5.17

Fig. 5.11 **(NB: many identical copies of a particular mRNA are present on blot)**

See also: Northerns of new ncRNAs (van der Meulen *RNA Biol*. 13: 353, 2016)

MOST COMMON TYPE of BLOT!

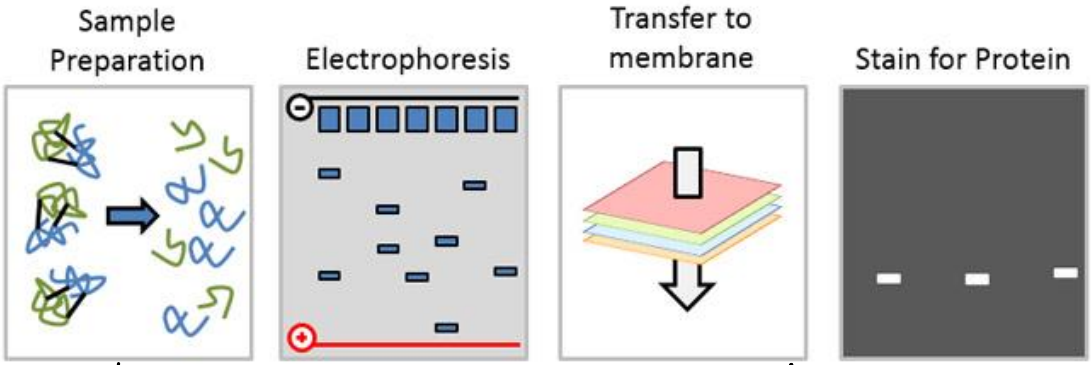
Developed in 1979

Other blots: Western Blot analysis (Immunoblotting)

- to identify proteins by probing with an antibody

– **Protein is electrophoresed, transferred to membrane and probed with an antibody**

- Is a specific protein present or not?
- Size of protein?
- abundance of protein?



Probe: primary antibody – *selective for protein of interest. Secondary antibody binds primary antibody and has aid for detection (eg horseradish peroxidase HRP)*

denature

SDS-PAGE

electrotransfer is used.

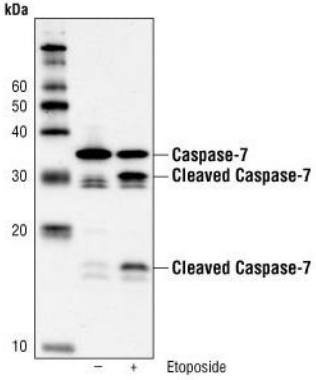
1° antibody - selective for target protein

2° antibody → detection.

HRP linked to 2°

oxidizes luminol

(ECL) enhanced chemiluminescence



Extracts of Jurkat cells treated with and without etoposide, an activator of apotosis, probed with anti-caspase-7 monoclonal antibody

See also: Global Histone Modification (Zhao *Clinical Epigenetics* 8:34, 2016)

Other Blots: Southwestern blot analysis

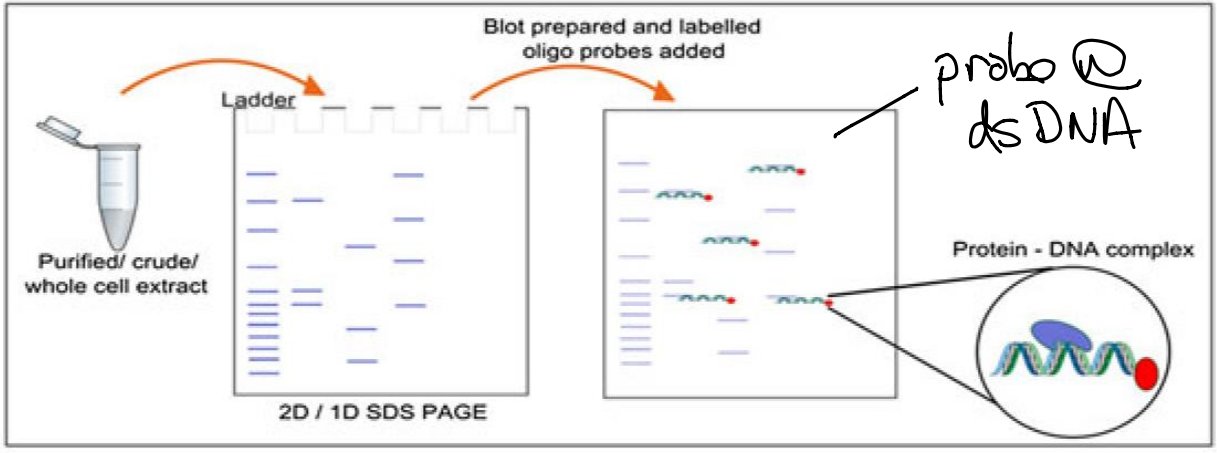
Developed in 1980

- to identify protein-DNA interactions

– **Protein is electrophoresed, transferred to membrane renatured and hybridized with DNA probe (eg promoter of interest)**

- Identify proteins that bind DNA
- Identify DNA sequence that binds protein

transcriptional regulation



probe @ dsDNA

Disadvantages: misses multi-subunit DNA binding proteins and proteins that do not refold

- modern expt
ChIP
Chromatin immunoprecipitation

Northwestern blot analysis

- to identify protein-RNA interactions

- probe @ RNA
- RNA binding proteins involved in post transcriptional regulation.
  ⇒ splicing, stability, localization & translation impacted by RNA binding prot.

Some protein encoding genes are constitutively expressed …

"housekeeping gene" products needed at all times

… whereas others are differentially expressed

in specific tissue type

during development

in response to environmental cues

transcriptional regulation - primary mechanism for regulating biological function

Only a subset of genes are expressed at a given time and mRNA levels can vary greatly among genes

~10,000 – 15,000 different mRNAs present in "typical" mammalian cell type under given condition

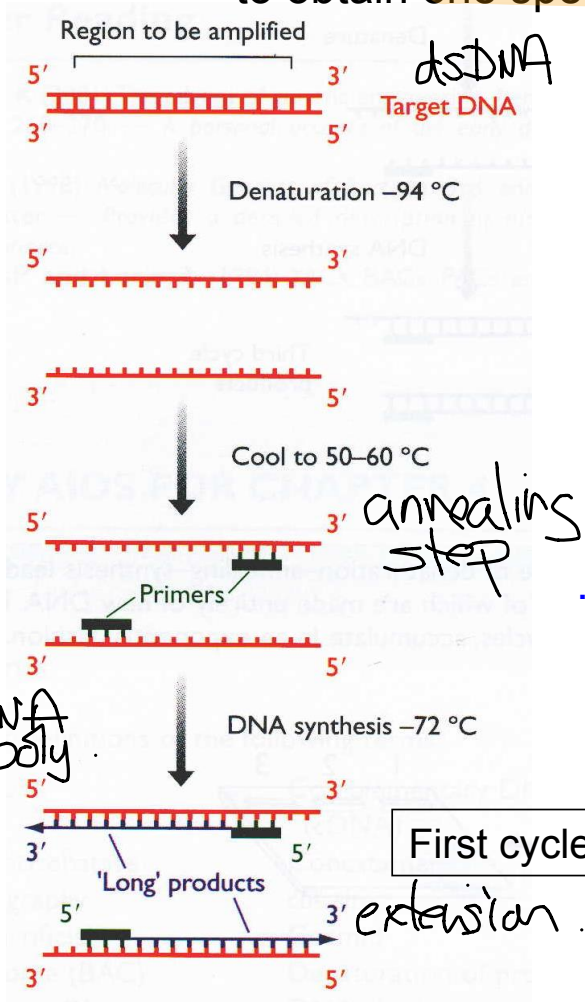
(~ 20,000 different proteins present)

Aside: RNA-sequencing studies suggest ~ 8000 genes ubiquitously expressed in human tissues (Ramskold *PLoS* 2009)

Developed in 1983 - takes advantage of DNA polymerase

3. PCR - polymerase chain reaction

- to obtain one specific DNA region in large copy number



dsDNA

annealing step

DNA poly.

First cycle

extension.

Fig.2.28

- rapid amplification of DNA region of interest by enzymatic reaction in test tube

1. Denaturation of duplex DNA

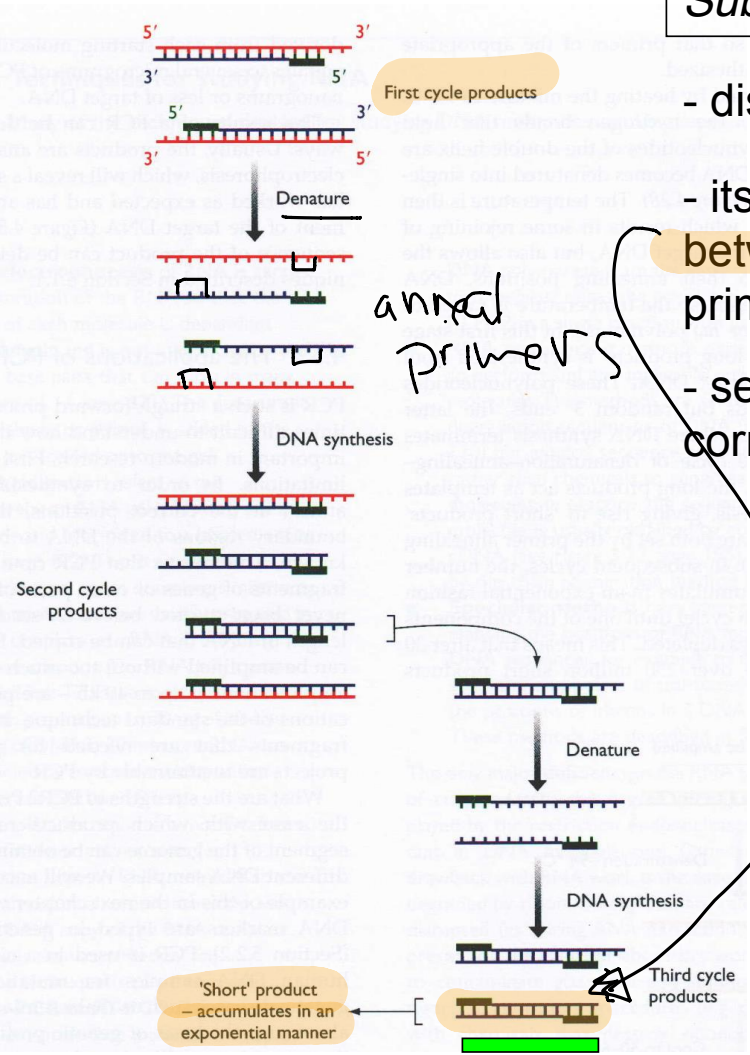
2. Annealing of 2 different primers
(synthetic oligomers, usually 15-25 nt)
- flank region of interest,
- in opposite orientation
… so anneal to opposite strands of DNA

3. Extension of complementary strands

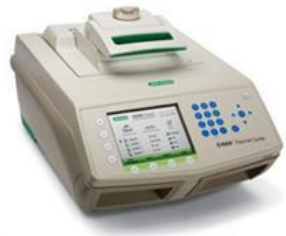
- cycle repeated 25-30 times

- polymerase synthesizes DNA from dNTPs in 5′–3′ direction.

Subsequent PCR cycles



- discrete PCR product generated
- its length corresponds to distance between primers (including the primers)
- sequences at ends of amplicon correspond to the 2 primers used



Bio-Rad C1000 Thermal Cycler

NB: even though text shows both primers as black bars, they have different nt sequences

Fig.2.29

Powers & pitfalls of PCR

- rapid method to generate large amounts of specific segment of DNA (product usually < 10 kb in length)

"Long-range PCR"
~ 20-30 kb products

- need very small amount of template DNA

…but can lead to contamination problems

- need prior sequence info to design primers

Typically use 20-25 nt oligomers,
but for simplicity (as on a test) 6'mers are shown here

Why choose ~ 20'mers?

Tip: see Question 2.5 in text (p.61)

Specificity!!!

A specific 20 mer sequence has less than 1% chance of occurring randomly in human genome.

probability of site $= \left(\frac{1}{4^{20}}\right) \times 3.2\,10^{9}$ bases $= 0.003$

$\simeq 0.3\%$

Essential to check PCR products!

How to double-check that PCR product (amplicon) is correct one?

1. Is it the right size?

- agarose gel electrophoresis (with size markers)

fast & easy always done

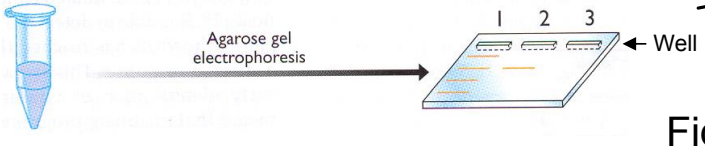
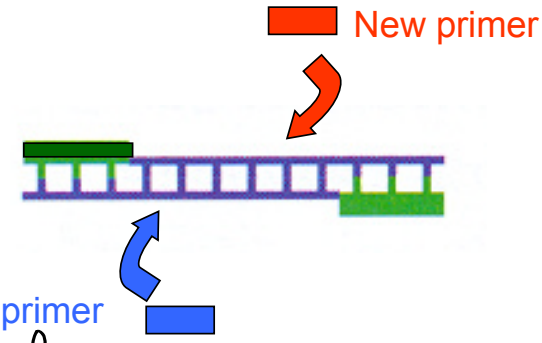


Fig. 2.30

2. Does it contain the right sequence (eg gene X)?

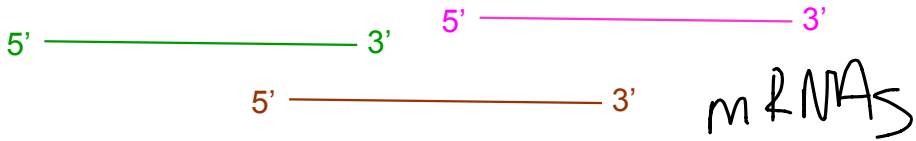
- Southern hybridization - using gene X (eg. clone) as probe
- restriction analysis - are expected restriction sites present?
- nested PCR
  - design "internal" primers to use in 2d PCR experiment with 1st PCR product as template DNA

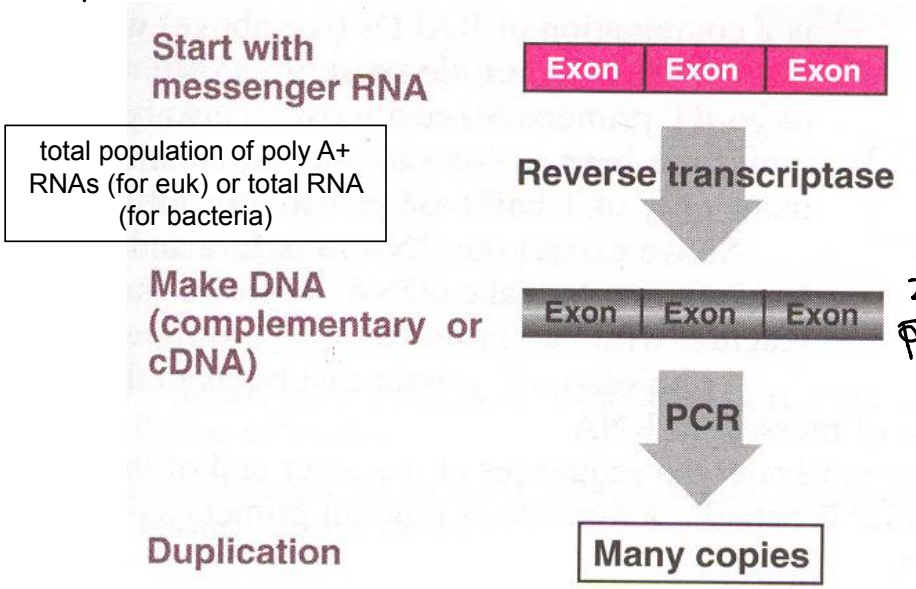


- Sequence PCR product most important method.

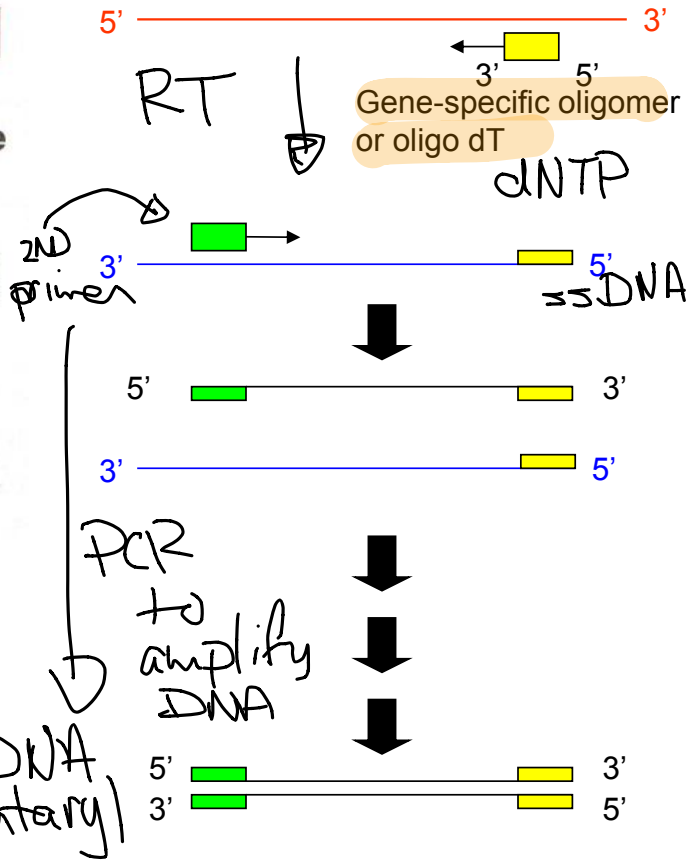
RT-PCR
Reverse Transcriptase

RNA to DNA





total population of poly A+ RNAs (for euk) or total RNA (for bacteria)



- then sequence RT- PCR product directly (or after cloning)

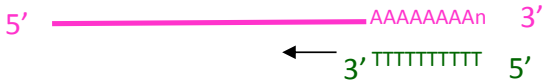
- (need sequence data to design primers for RT-PCR)

RT come from RNA viruses

(see p.142, Chapter 5)

Choice of primer for cDNA synthesis

1. oligo dT - anneals to polyA tail of mRNA (if eukaryotic)

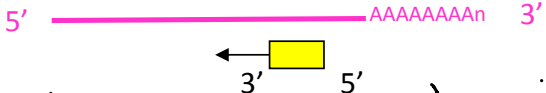


Generates copies of all mRNAs present

2. Random short primers (eg. 6'mers)

Generates partial copies of mRNAs

3. Sequence-specific primer

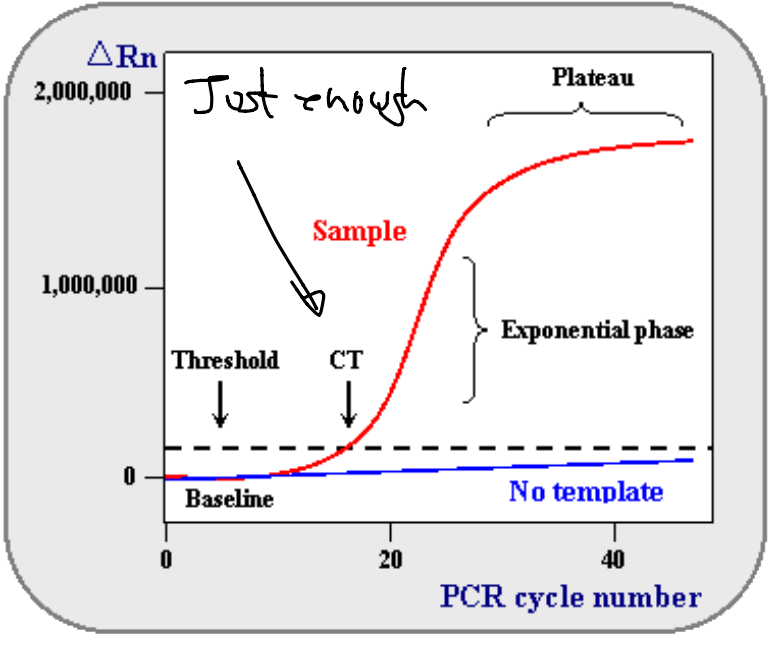


Generates cDNA just for specific gene of interest

- disadvantage of poly(T) primer - large 3′ UTR (>1Kb) may not get 5′ end of mRNA
- gene specific primer - very good for rare mRNA

Real-time quantitative PCR (qPCR)

- to measure relative or absolute amount of mRNA present in different tissue types/developmental stages/environmental conditions…



ΔRn : increment of fluorescent signal at each time point

CT : PCR cycle number where reporter fluorescence is greater than threshold

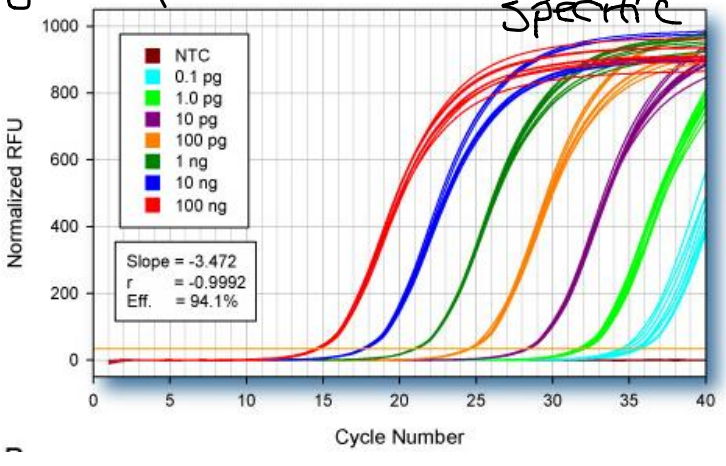
- detection and measurement of products generated during each cycle of PCR by using a reporter fluorescent probe

2 types of detection

eg. SYBR green, TaqMan

general

segment specific



RFU = relative fluorescence units
NTC = no template control

http://www.ncbi.nlm.nih.gov/projects/genome/probe/doc/TechQPCR.shtml
NCBI Technologies website

4. CLONING - to obtain one specific DNA region in large copy number

- by using host cell (eg. E.coli) to amplify DNA of interest

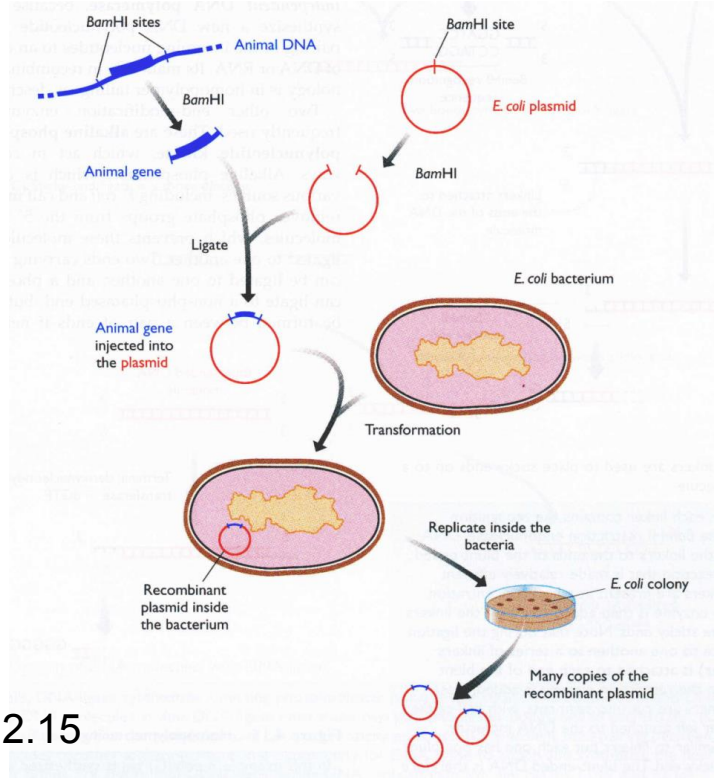


Fig. 2.15

- DNA fragments ligated into vector

… then introduced into bacterial (or yeast…) cell

by transformation

to generate clone library

= collection of clones whose inserts cover the entire genome

Each colony has many identical copies of recombinant plasmid & independent colonies contain different regions of genome

Number of clones needed for library to cover complete genome will depend on genome size & insert size in vector (Table 2.4)

Aside: cDNA library - mRNAs reverse-transcribed into cDNAs and cloned (Fig. 5.32)

Examples of cloning vectors used to generate clone library

1. Plasmid

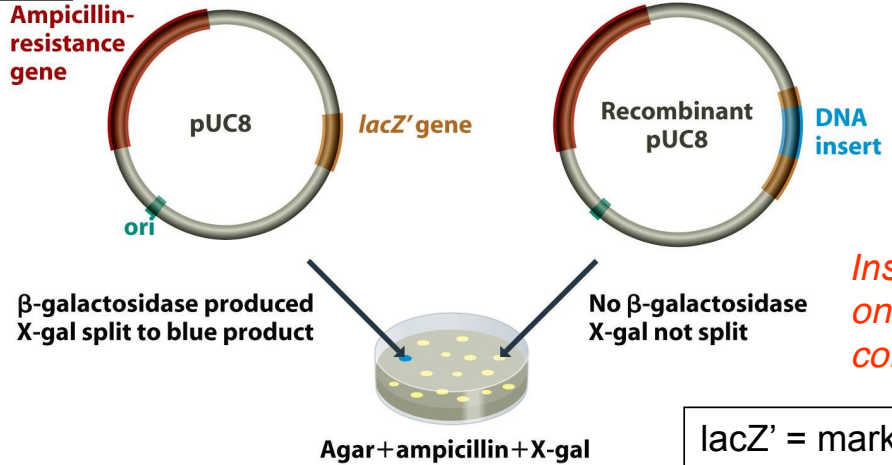
Table 2.4

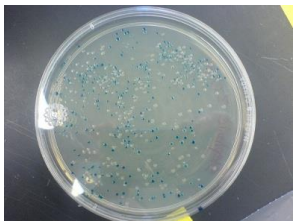
- to clone < 10 kb fragments
- origin of replication, selectable markers

eg. antibiotic resistance in bacteria: ampicillin, tetracycline …

or nutrient requirement in yeast: URA3, TRP1 …

Selectable marker





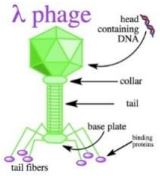
Insert disrupts lacZ' gene, so Xgal on plate not converted to blue colour & colonies are white

lacZ' = marker for rapid screening of recombinants

Fig.2.18

2. Phage lambda

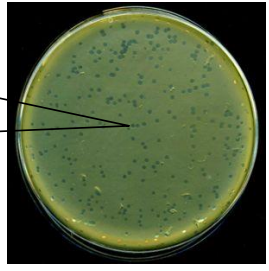
- to clone 15-20 kb DNA fragments



Mid-region of λ DNA molecule can be removed and replaced with similar-sized insert DNA of interest, then packaged in phage particle

(Aside: also λ vectors for cloning cDNAs of 1-5 kb)

Plaques (ie. clearing on lawn of E.coli where phage have lysed the bacteria)



3. Cosmid

- λ-plasmid hybrid, cos site to package DNA in phage particle
- to clone ~40-45 kb fragments
- fosmid: low copy number (more stable) cosmid vector based on E.coli F-factor replicon

- popular for microbial metagenomics projects

4. BAC

- bacterial artificial chromosome (~8 kb) with F (fertility) plasmid origin of replication
- to clone ~ 300 kb fragments

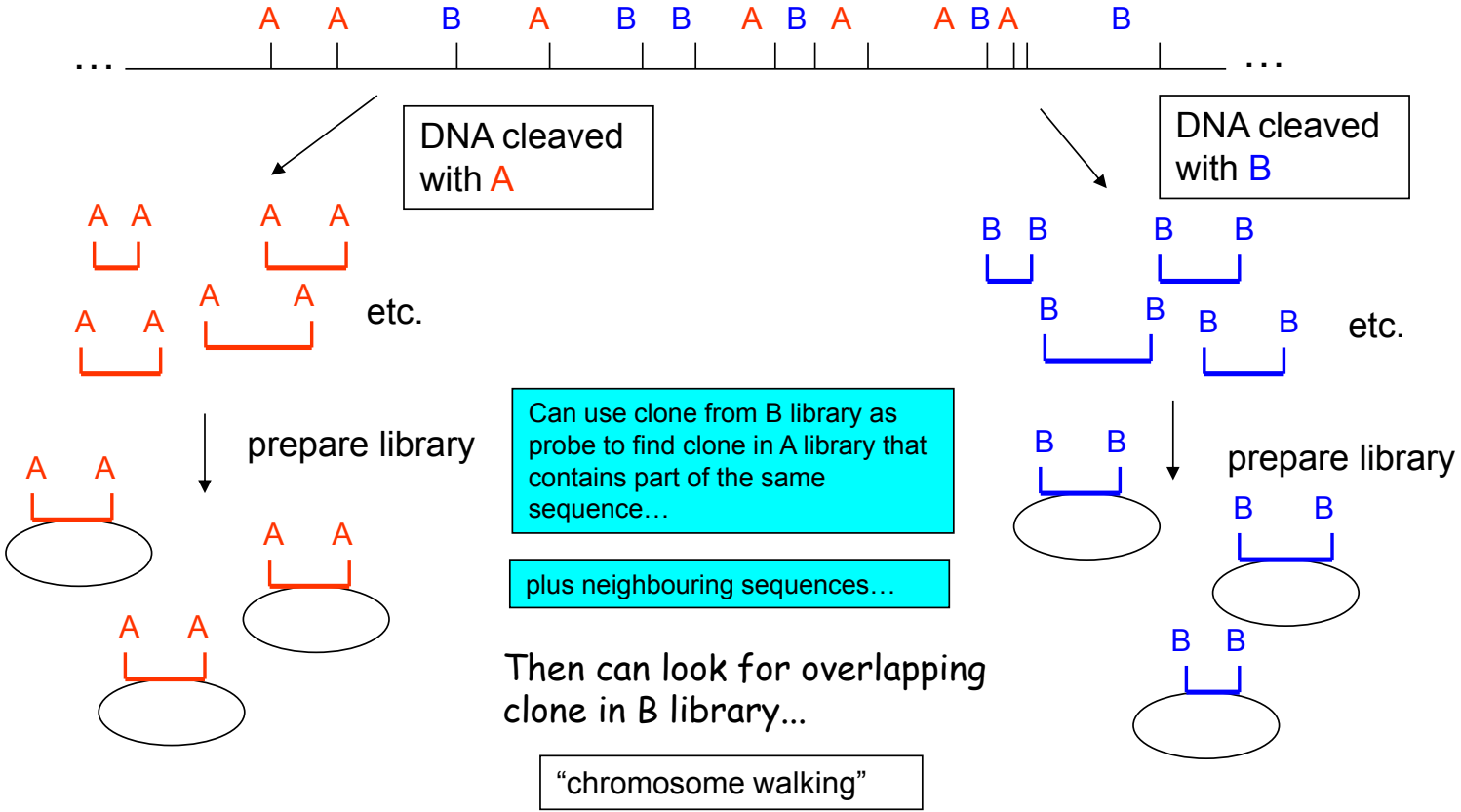
- most commonly used vector for cloning large DNA fragments

5. YAC

- yeast artificial chromosome (Fig.2.25)
- to clone ~ 1 Mbp fragments but inserts often unstable (DNA rearrangements)

Strategies for genomic clone library formation (for genome map assembly)

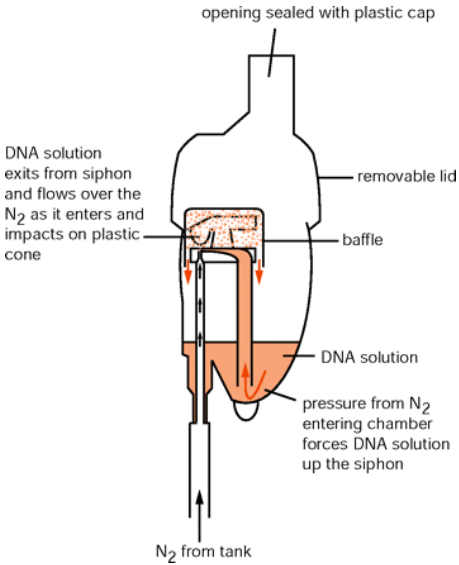
1. Use two clone libraries with restriction fragments derived from *different* restriction enzymes (or from incomplete digestion with one restriction enzyme)



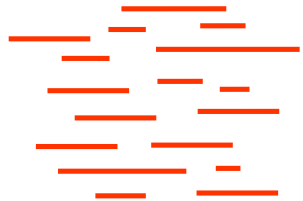
2. Random fragmentation of DNA (eg sonication or nebulization)

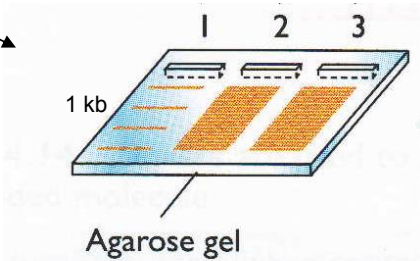
then blunt-end ligation into vector (or ligation into vector after *linkers* containing restriction sites added)

(see Fig. 2.13 for discussion of linkers)

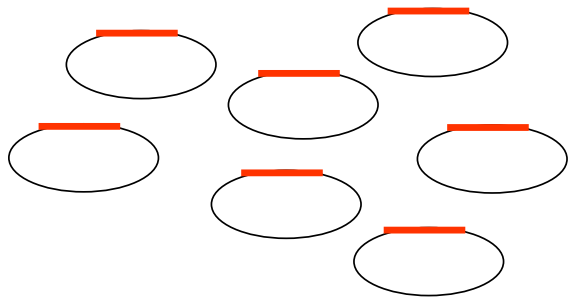


Nebulizer for random shearing of DNA





Recover DNA of desired size (eg. 1 kb) by gel electrophoresis (or repeated nebulization) & prepare clone library



... having random overlapping segments of genome

Cold Spring Harbor Protocols 2010

Aside: this method was used to obtain the first complete bacterial genome sequence (*Haemophilus influenza*) (Topic 5 & Fig.4.10)