

# **Analysis of Gene Expression**

**SC113603 Molecular Biology**  
Lect. Todsapol Techo

# **Objective**

- Describe about the techniques of DNA technology related to the measurements of gene expression
- Describe about the techniques of protein analysis related to gene expression

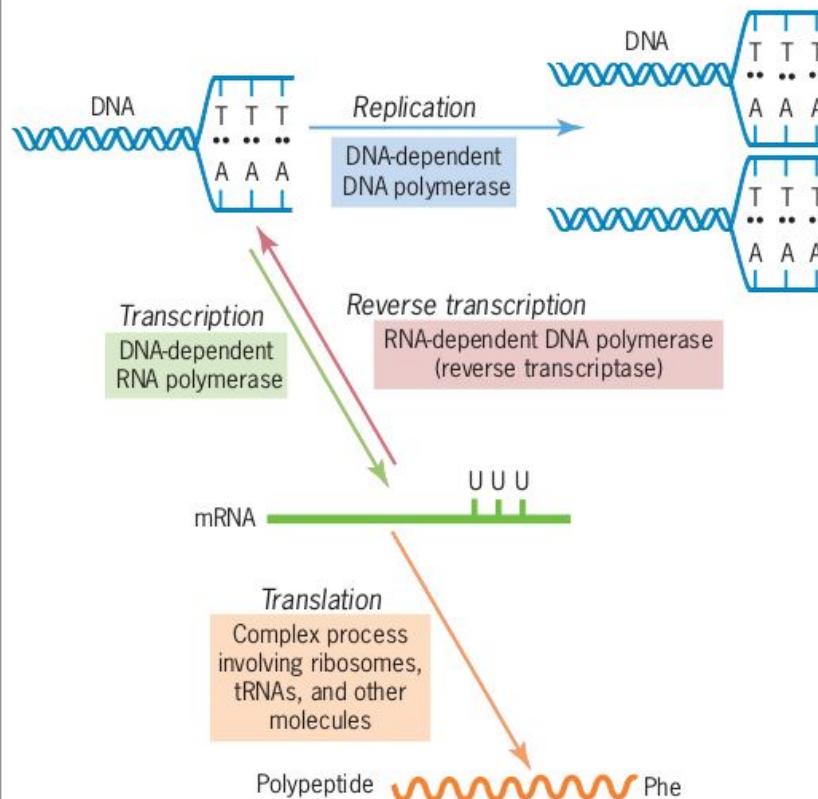
## The Central Dogma

Flow of genetic information:

1. Perpetuation of genetic information from generation to generation

2.  
Control  
of the  
phenotype:

Gene  
expression



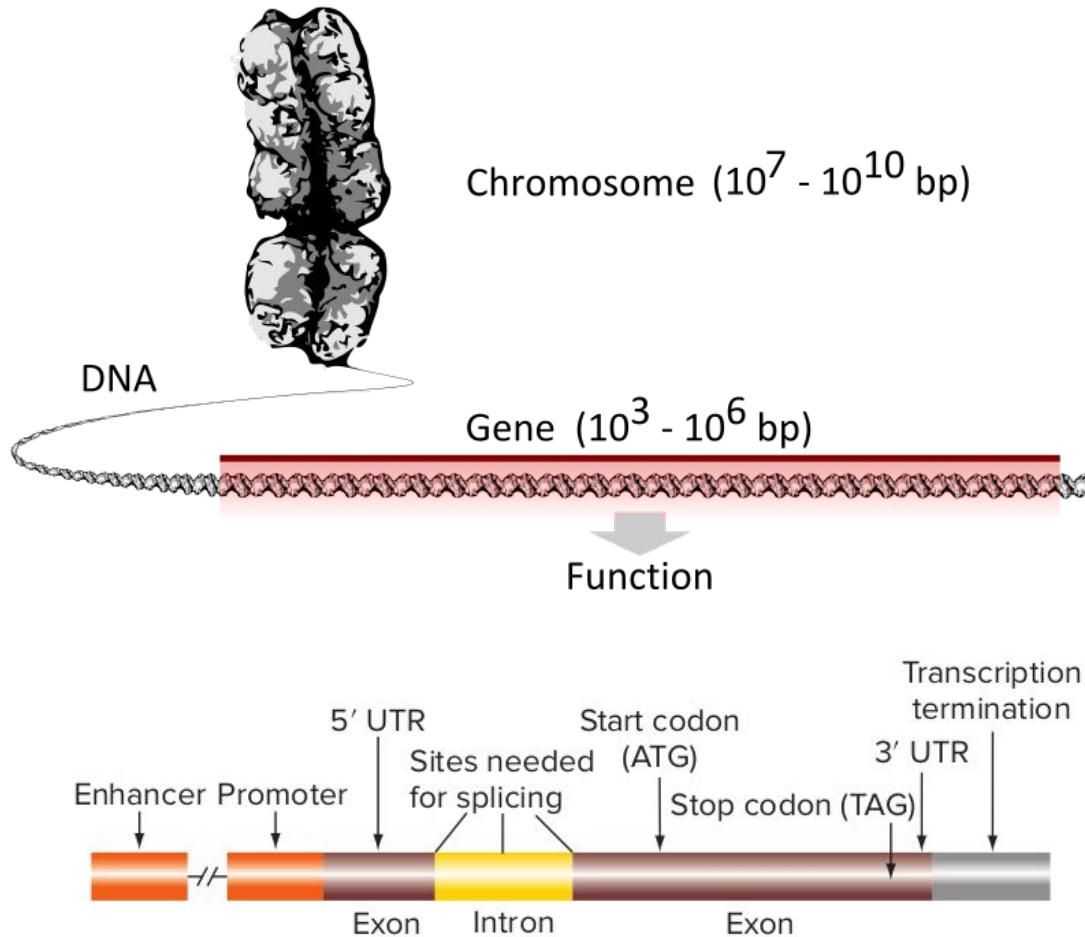
# Gene Expression

- **Central Dogma in Molecular Genetics**

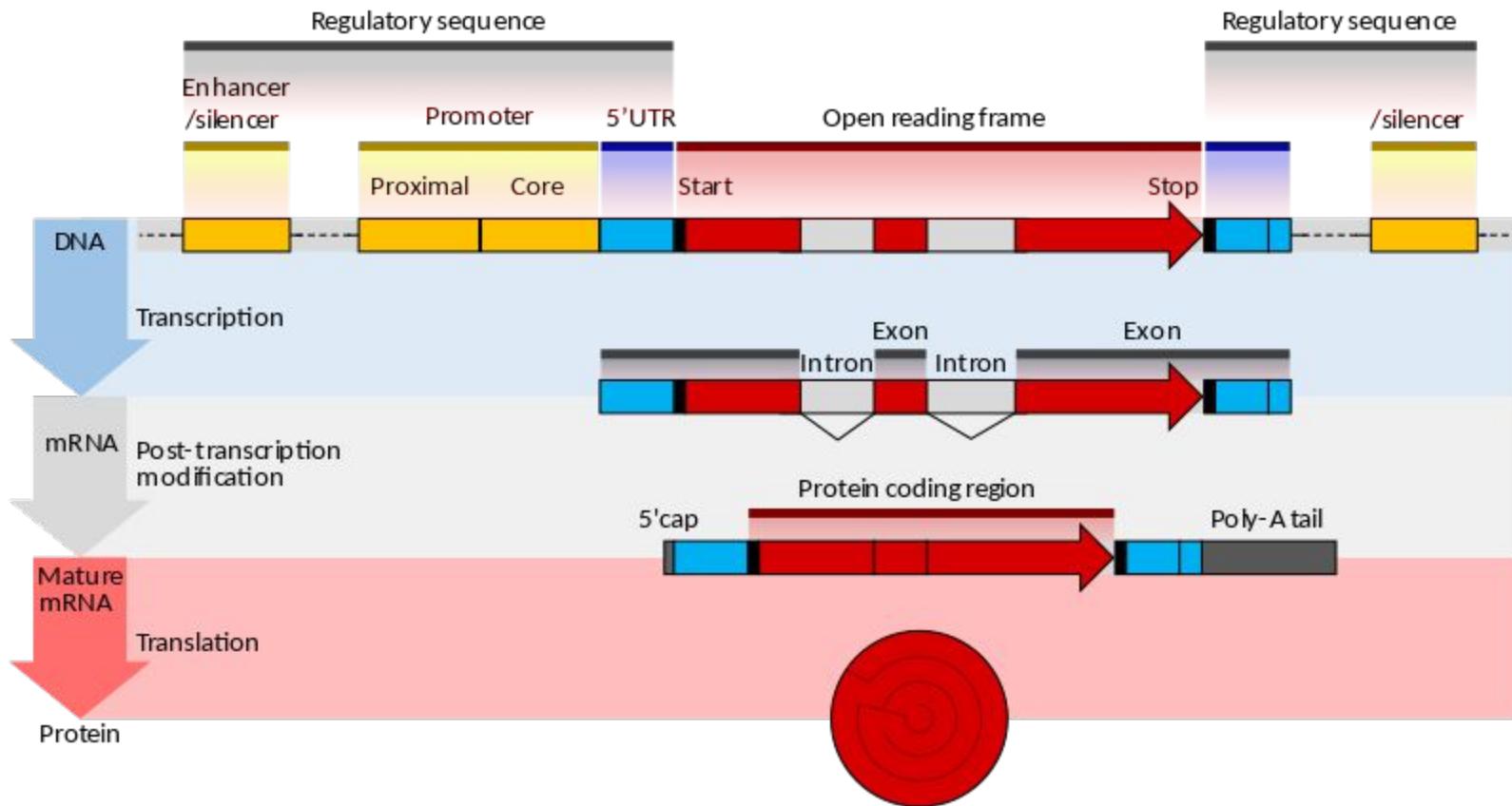
- **Replication**
  - Synthesis of Genetic material
- **Transcription**
  - Copy DNA to mRNA
- **Reverse Transcription (In Virus)**
  - Copy RNA to DNA
- **Translation**
  - Translate mRNA to Polypeptide (Protein)

# Gene

- **Specific region of DNA sequences that encode a function**
  - Could be **transcribed** to mRNA
  - Could be **translated** to Protein
  - Affect Phenotypic Traits
- **Composed of Two main Regions**
  - **Coding region**
    - Decode to Protein
  - **Noncoding region**
    - Regulate The flow of genetic information in coding region



# Gene Structure in Eukaryotes



# Measurements of Gene Expression

- RNA-based Measurements

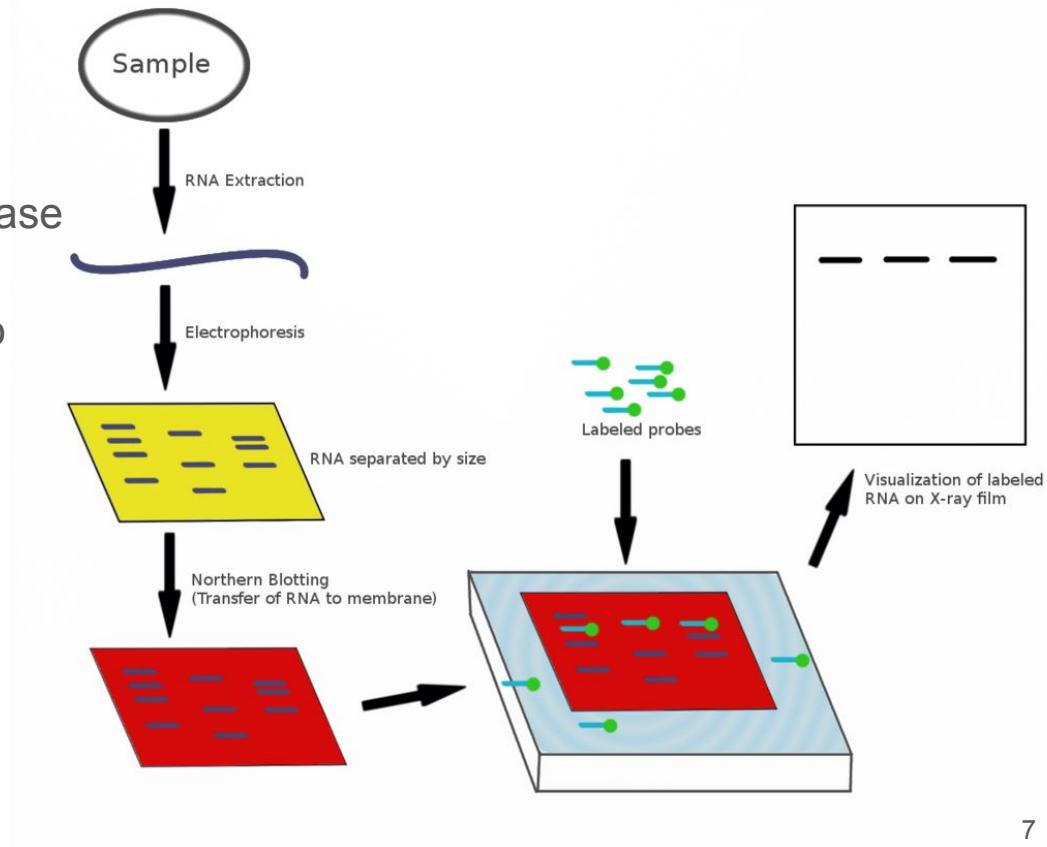
- Low-Throughput
  - Northern Blot: Based on Hybridization technique and fluorescent probe
  - RT-PCR: Based on PCR and Reverse Transcriptase (RT)
  - RT-qPCR: Based on PCR with fluorescent dye
- High-Throughput
  - Microarray: Based on Hybridization technique and fluorescent probe
  - RNA-seq: Based on NGS

- Protein-based Measurements

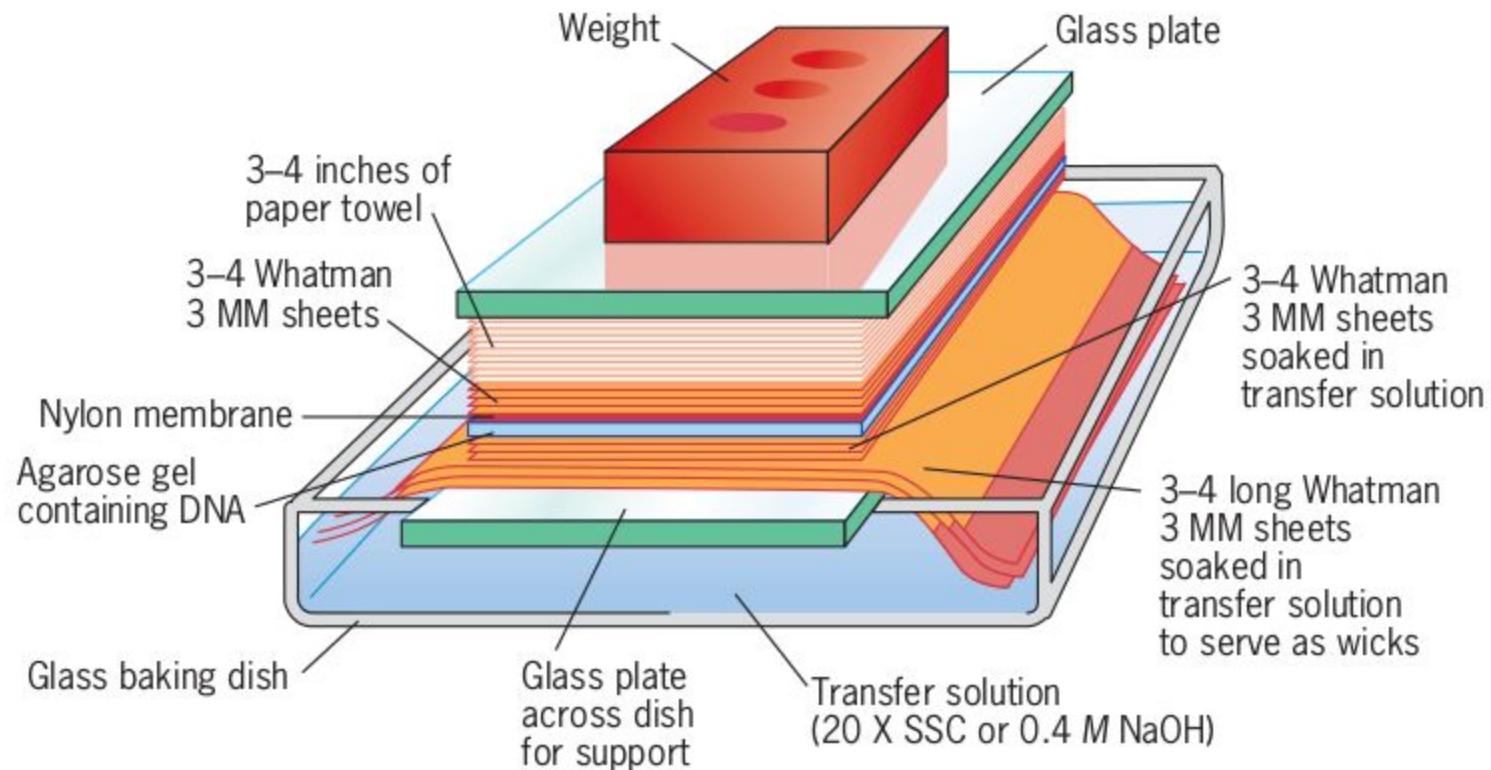
- Low-Throughput
  - Western Blot: Based on antigen-antibody binding and PAGE
- High-Throughput
  - Mass Spectrometry
    - Two-Dimensional Polyacrylamide Gel Electrophoresis (2D-PAGE) -> MS
    - LC and HPLC -> MS

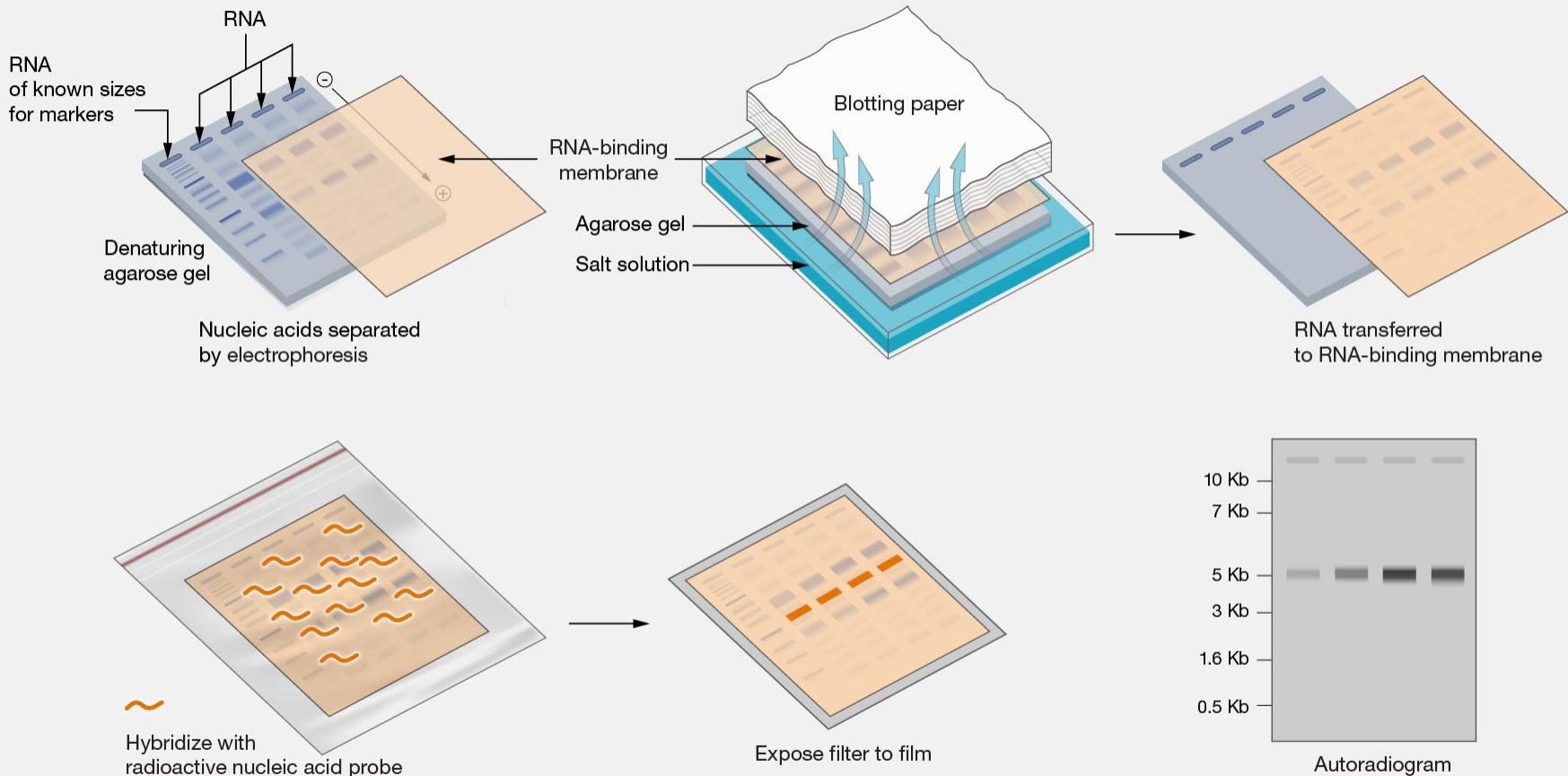
# Analysis of RNA by Northern Blot

- **RNA Isolation:** NO RNase Contamination
- **RNA gel electrophoresis:** NO RNase Contamination
- **Blotting:** Transfer RNA from Gel to RNA-binding Membrane
- **Hybridization:**  
Fluorescent/Radioactive Probe
- **Detection:** By using fluorescent detector or film

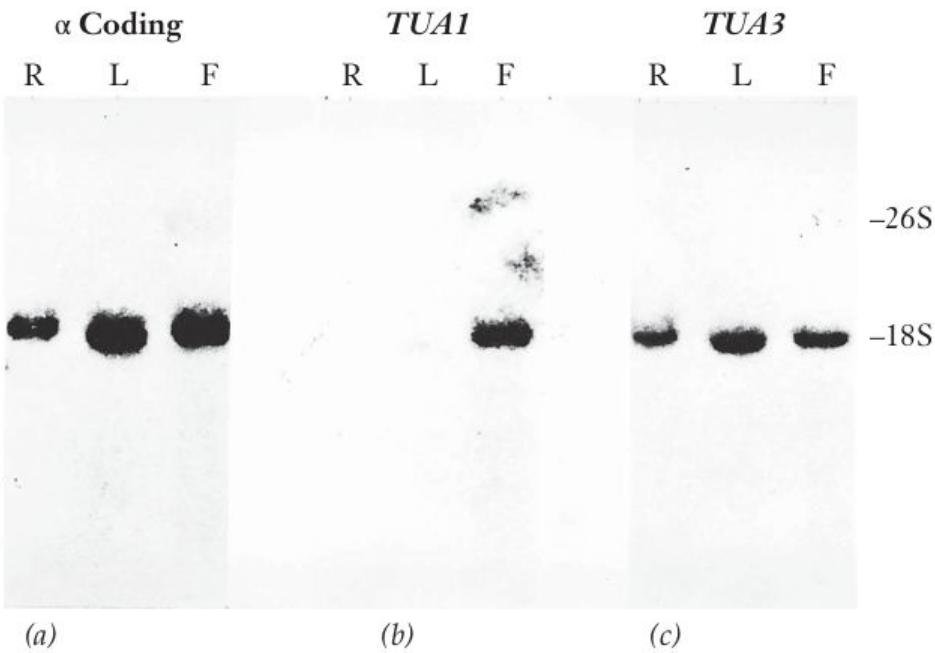


# Blotting: Transfer DNAs from Gel to membrane





Courtesy of S. R. Ludwig and D. P. Snustad,  
University of Minnesota. Page 383; From Kerem,  
et al. (1989). Science 245: 1073–1080.



**FIGURE 14.13** Typical northern blot hybridization data. Total RNAs were isolated from roots (R), leaves (L), and flowers (F) of *A. thaliana* plants, separated by agarose gel electrophoresis, and then transferred to nylon membranes. The autoradiogram shown in (a) is of a blot that was hybridized to a radioactive probe containing an  $\alpha$ -tubulin coding sequence. This probe hybridizes to the transcripts of all six  $\alpha$ -tubulin genes in *A. thaliana*. The autoradiograms shown in (b) and (c) are of RNA blots that were hybridized to DNA probes specific for the  $\alpha 1$ - and  $\alpha 3$ -tubulin genes (*TUA1* and *TUA3*, respectively). The results show that the  $\alpha 3$ -tubulin transcript is present in all organs analyzed, whereas the  $\alpha 1$ -tubulin transcript is present only in flowers. The 18S and 26S ribosomal RNAs provide size markers. Their positions were determined from a photograph of the ethidium-bromide stained gel prior to transfer of the RNAs to the nylon membrane.

# Measurements of Gene Expression

- RNA-based Measurements

- Low-Throughput
  - Northern Blot: Based on Hybridization technique and fluorescent probe
  - RT-PCR: Based on PCR and Reverse Transcriptase (RT)
  - RT-qPCR: Based on PCR with fluorescent dye
- High-Throughput
  - Microarray: Based on Hybridization technique and fluorescent probe
  - RNA-seq: Based on NGS

- Protein-based Measurements

- Low-Throughput
  - Western Blot: Based on antigen-antibody binding and PAGE
- High-Throughput
  - Mass Spectrometry
    - Two-Dimensional Polyacrylamide Gel Electrophoresis (2D-PAGE) -> MS
    - LC and HPLC -> MS

# PCR

## Repeat Reaction

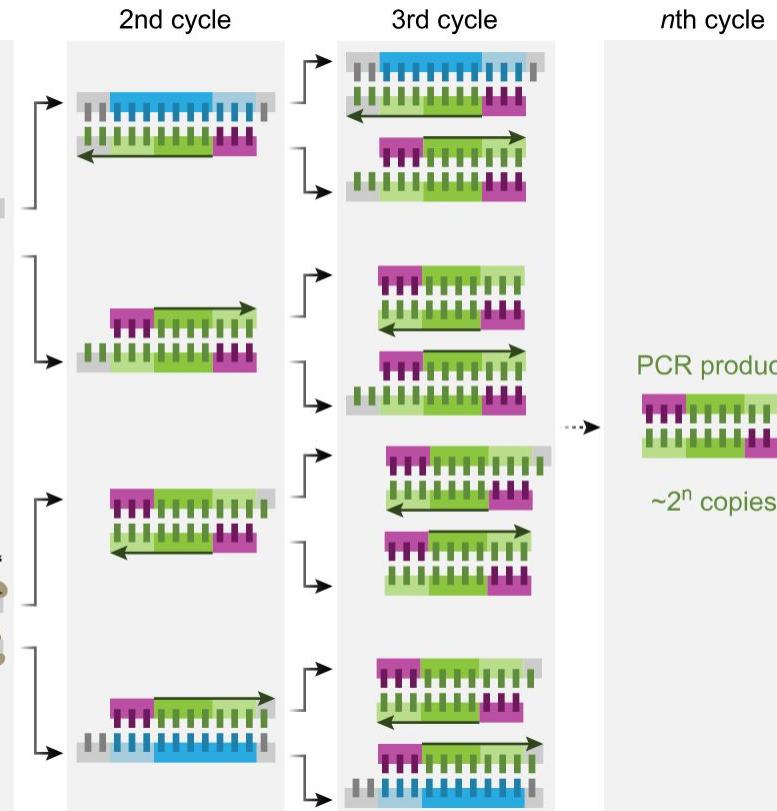
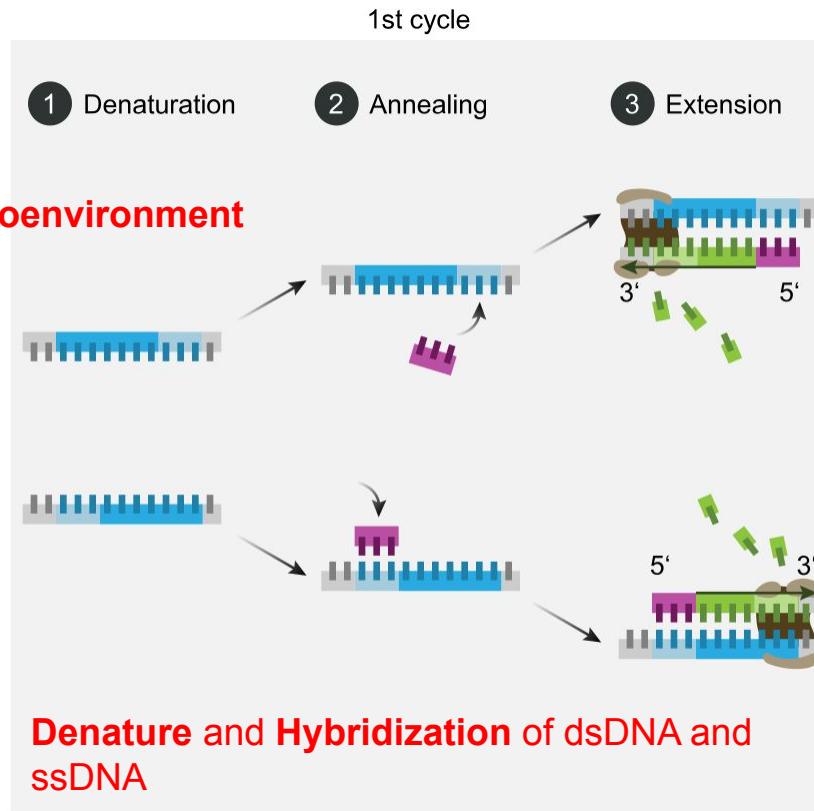
### Simulate Microenvironment

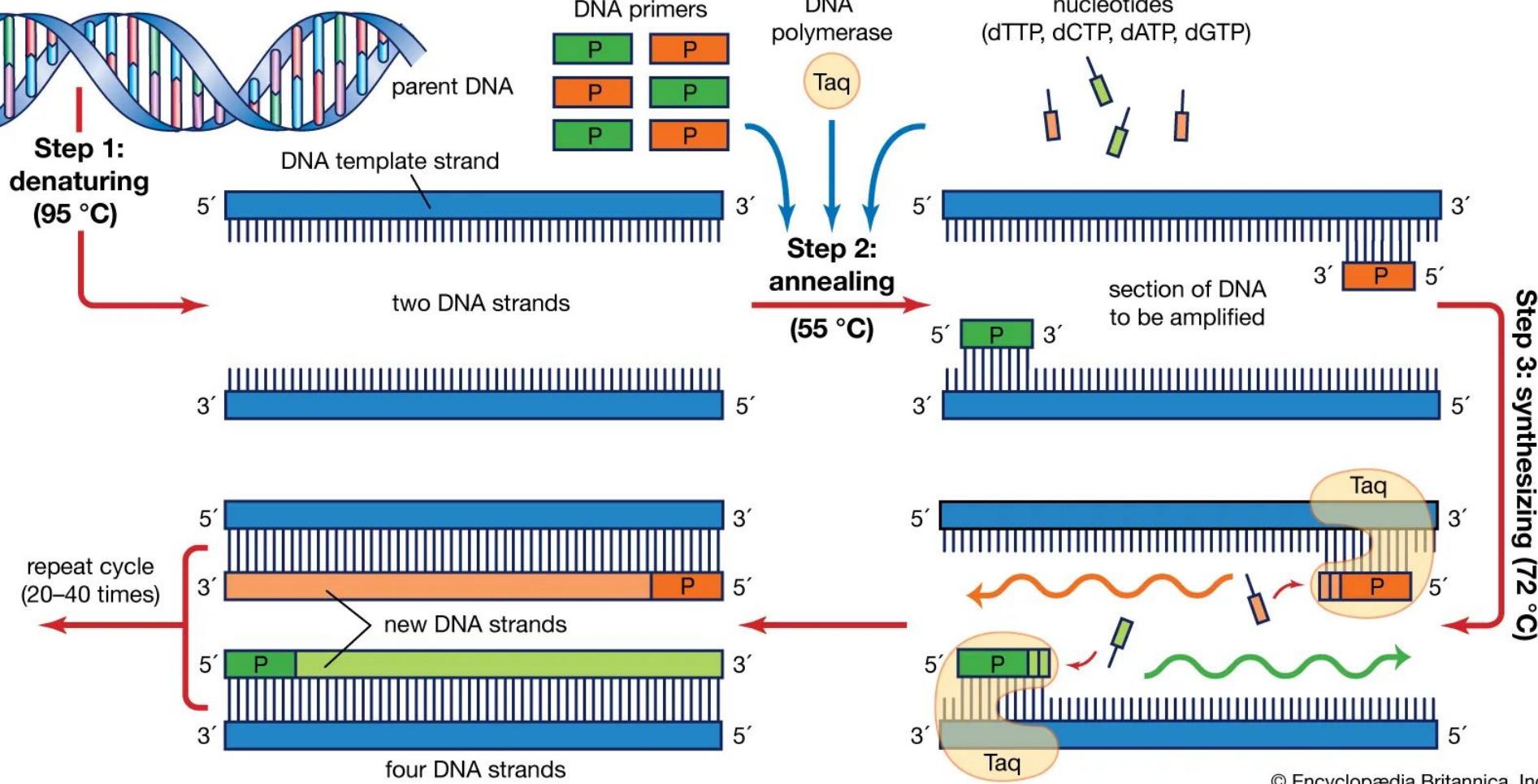
DNA template with sequence of interest  
5' 3'  
3' 5'

dNTPs

Primers

Polymerase

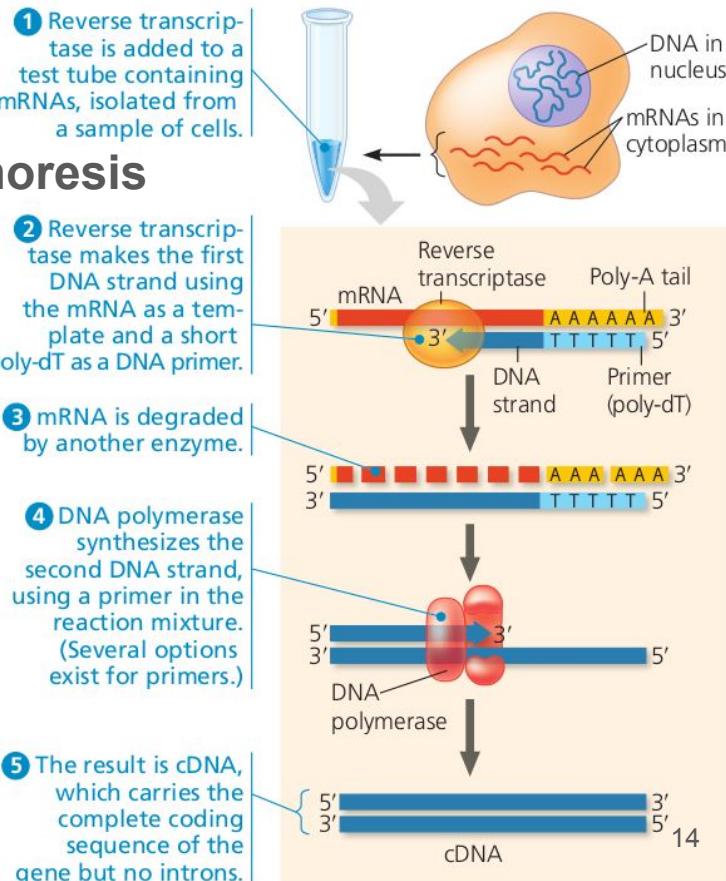




# Reverse Transcriptase (RT) PCR

- RT-PCR
- Use PCR technique with RT and gel electrophoresis
  - Use enzyme **RT** to convert RNA to **complementary DNA (cDNA)**
  - Use primer to specific amplify target cDNA
  - Use gel electrophoresis to separate fragments by size
- Reverse Transcriptase
  - Found in RNA Virus
  - Reverse transcription
    - Convert mRNA (ssRNA) to cDNA (dsDNA)
  - Required **Primer (poly-dT)**

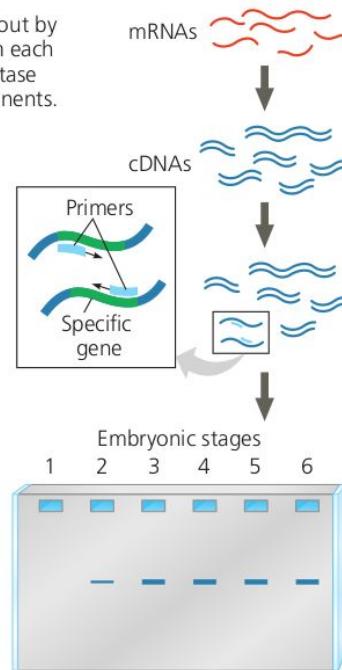
▼ **Figure 20.10 Making complementary DNA (cDNA) from eukaryotic genes.** Complementary DNA is made in a test tube using mRNA as a template for the first strand. Only one mRNA is shown after step 1, but the final collection of cDNAs would reflect all the mRNAs present in the cell.



## Reverse transcription-PCR

# RT-PCR

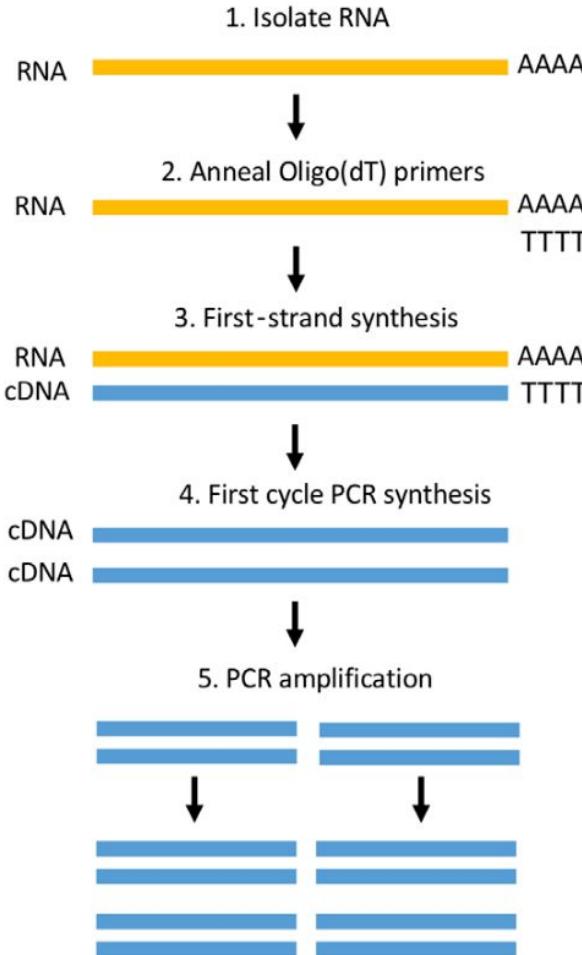
1 **cDNA synthesis** is carried out by incubating the mRNAs from each stage with reverse transcriptase and other necessary components.



2 **PCR amplification** is then performed using primers specific to the *Drosophila* gene of interest to see whether its mRNA was present in each sample.

# PCR

3 **Gel electrophoresis** will reveal amplified DNA products only in samples that contained mRNA transcribed from the specific *Drosophila* gene.

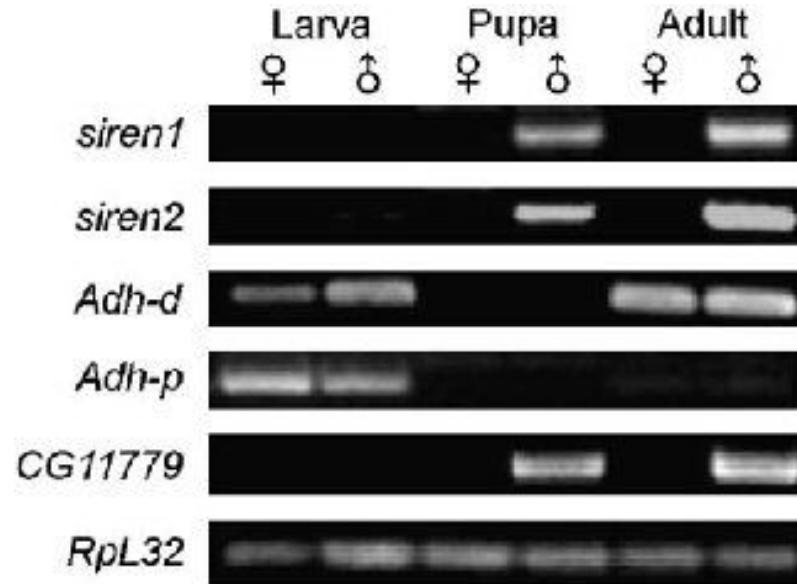


# RT

# > Gel Ele.

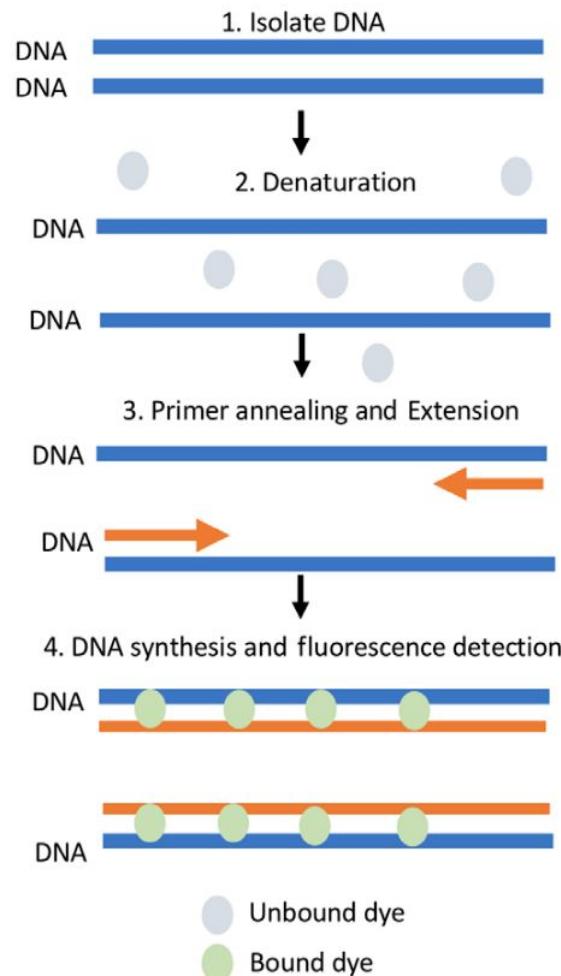
# RT-PCR

- Example of RT-PCR gel electrophoresis
  - Required **Reference Gene**: Gene that expressed all the time (*RpL32*)



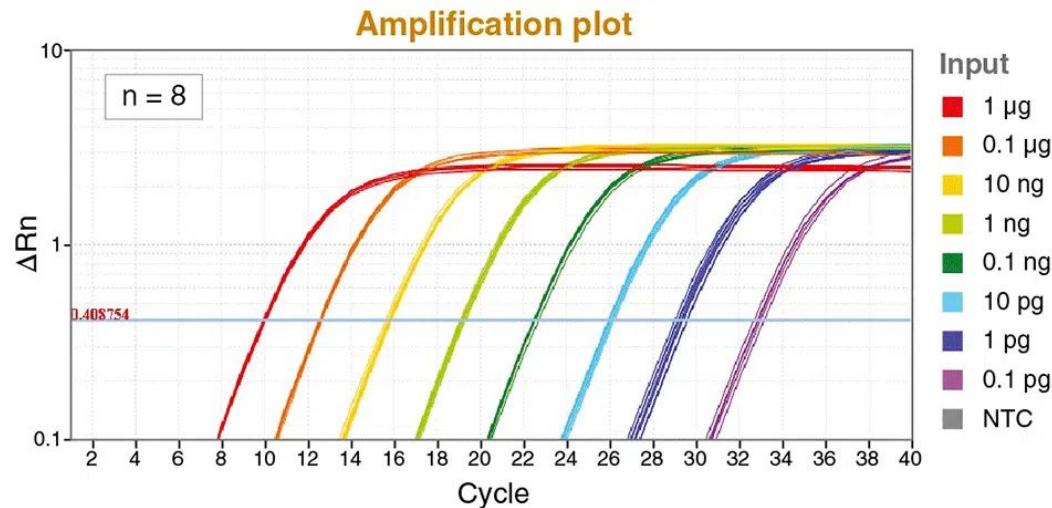
# Quantitative real time-PCR (qPCR)

- Based on PCR
- **Fluorescent Dye** Bind to **dsDNA**
- **Fluorescent Probe** emit signal in extension step
- The Quantitative measurement of DNA or cDNA
- Use qPCR instruments to detect fluorescent intensity



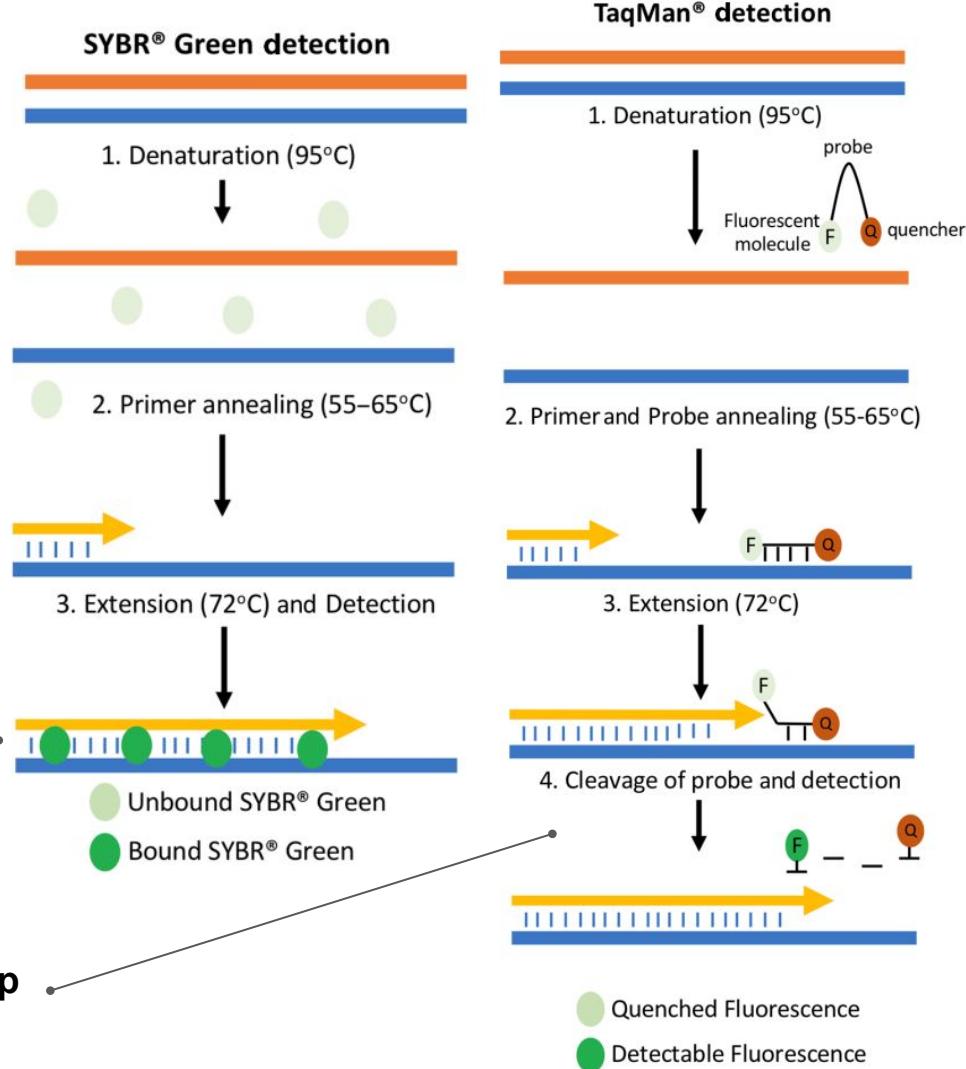
# qPCR Result

- S1 DNA 1 → 8 (3 cycles)
- S2 DNA 2 → 16 (3 cycles)
- S3 DNA 3 → 24 (3 cycles)
- S4 DNA 10 → 80 (3 cycles)
- S5 DNA 100 → 800 (3 cycles)
- S6 DNA 0 → 0 (3 cycles)
- If we can detect when DNA fragments more than 10 fragments
- Cycle 1 → S4, S5
- Cycle 2 → S3, S4, S5
- Cycle 3 → S2, S3, S4, S5



# Systems of qPCR

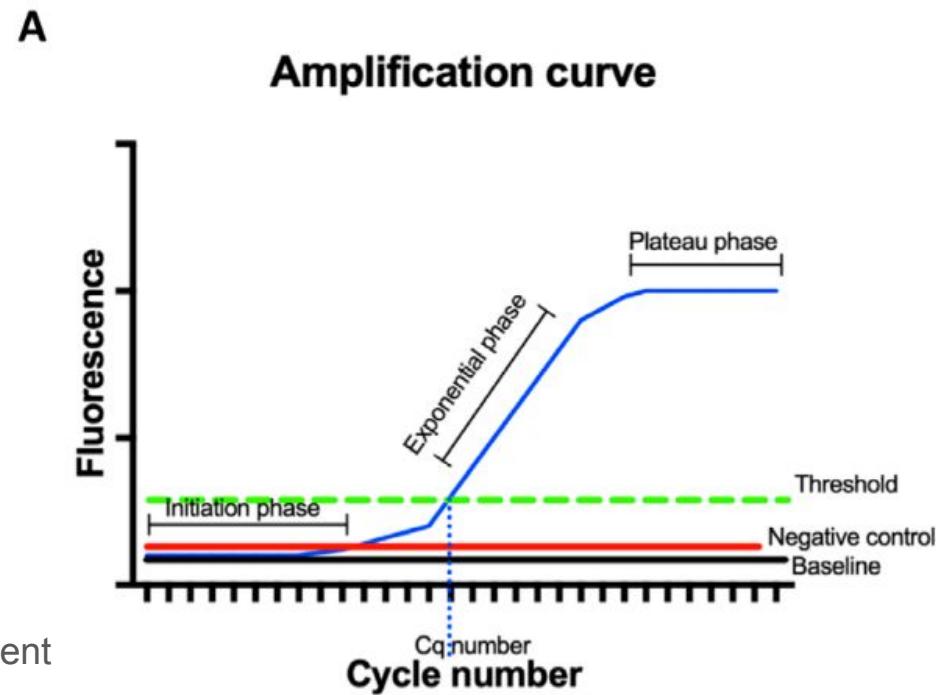
- Fluorescent Dye
  - Bind only double strands DNA
- Fluorescent Probe
  - Quencher: Block Fluorescent



Fluorescent Detection in Extension step

# qPCR Result

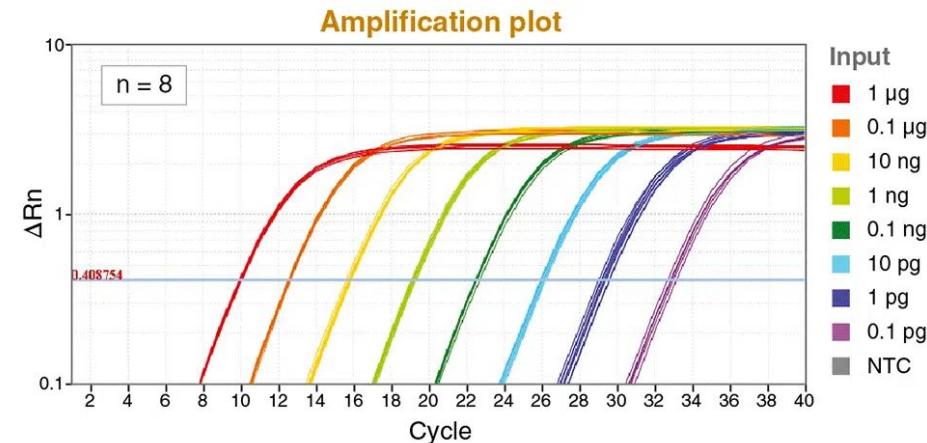
- Y-axis: Fluorescent Intensity
- X-axis: Cycle number
- Initiation Phase
  - Low Fluorescent Signal
  - Low DNA fragment
- Exponential Phase (High Signal)
  - Continuous increasing of Signal
  - Continuous increasing of DNA fragment
- Plateau Phase (Saturated Signal)
  - Saturated Fluorescent signal
  - May Continuous increasing of DNA fragment



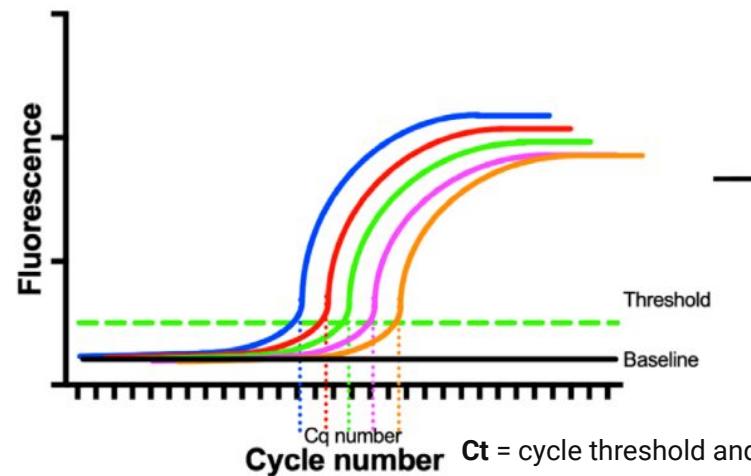
**C<sub>t</sub>** = cycle threshold and **C<sub>q</sub>** = quantification cycle

# Quantitative measurement of DNA

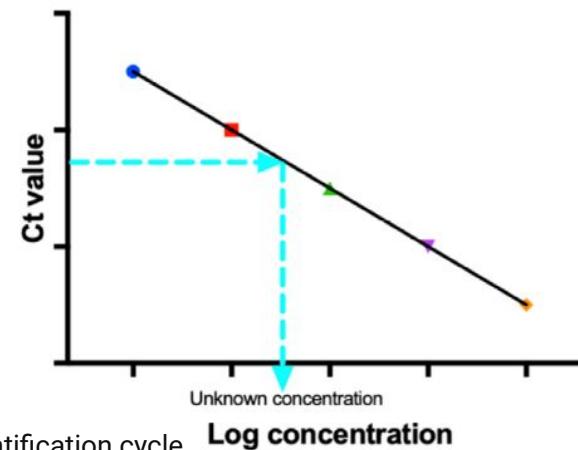
- qPCR
- Standard Curve (Regression Equation)



B Amplification standard curve

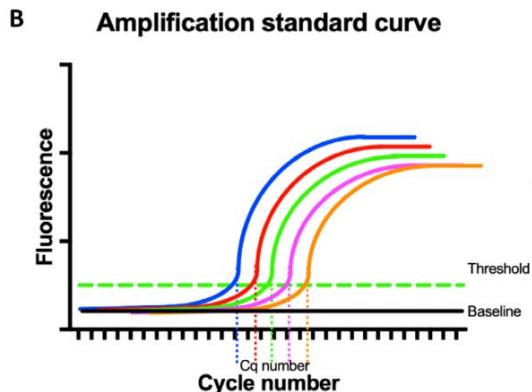


Standard Curve of Dilutions

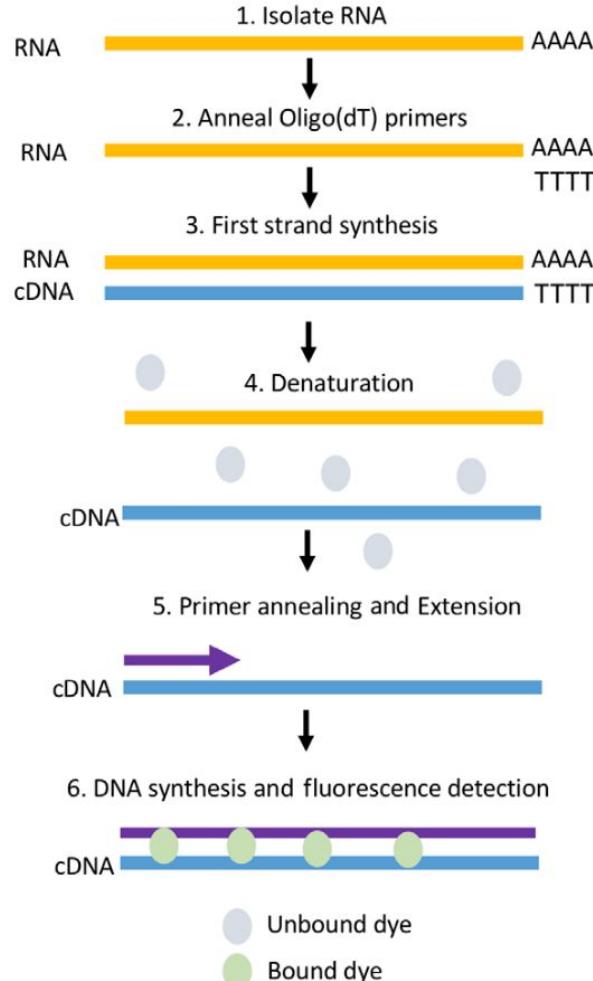


# RT-qPCR

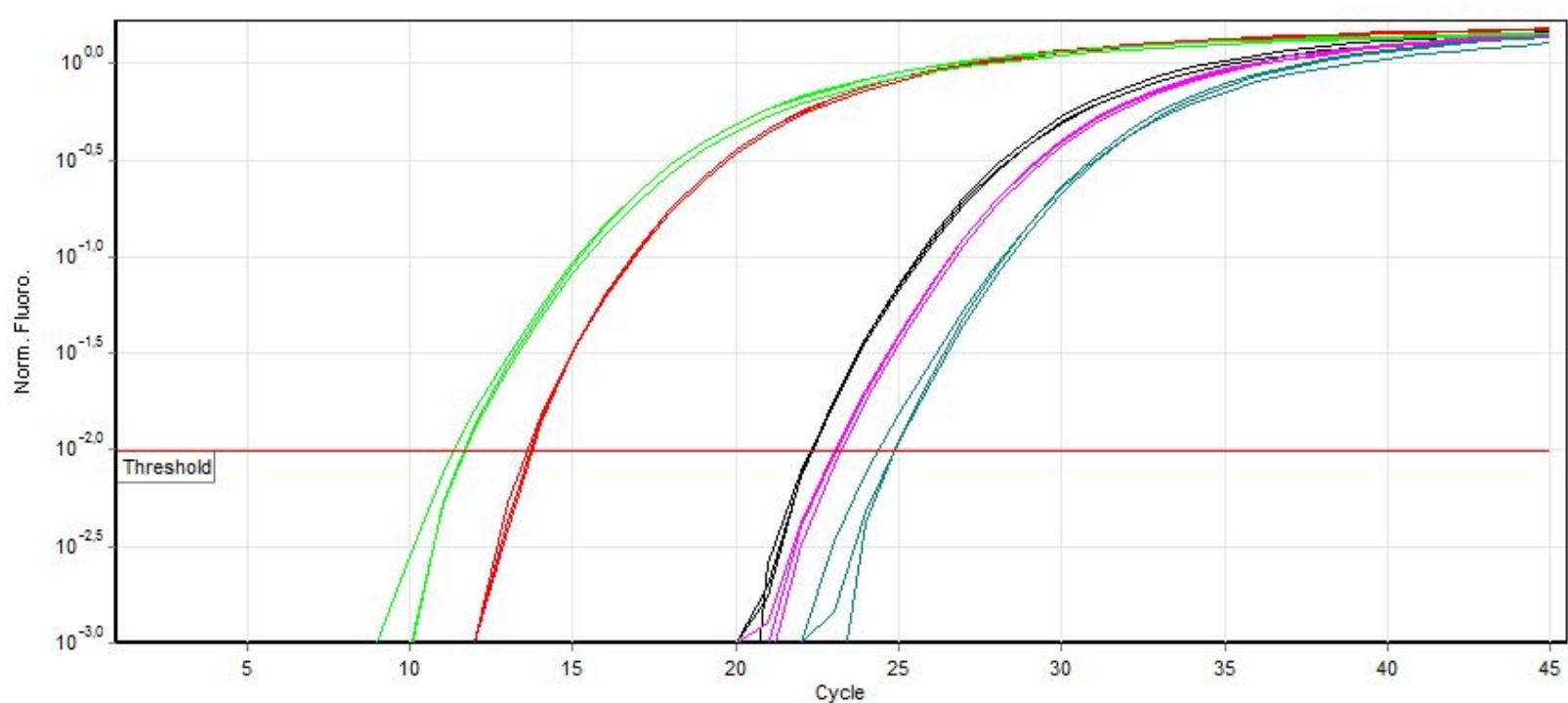
- Required Specific PCR primer for specific genes
- RT > convert mRNA to cDNA
- PCR to enrich target cDNA
- Fluorescent dye or probe to detect specific fragment in each PCR cycle
- Amplification curve is generated as the result

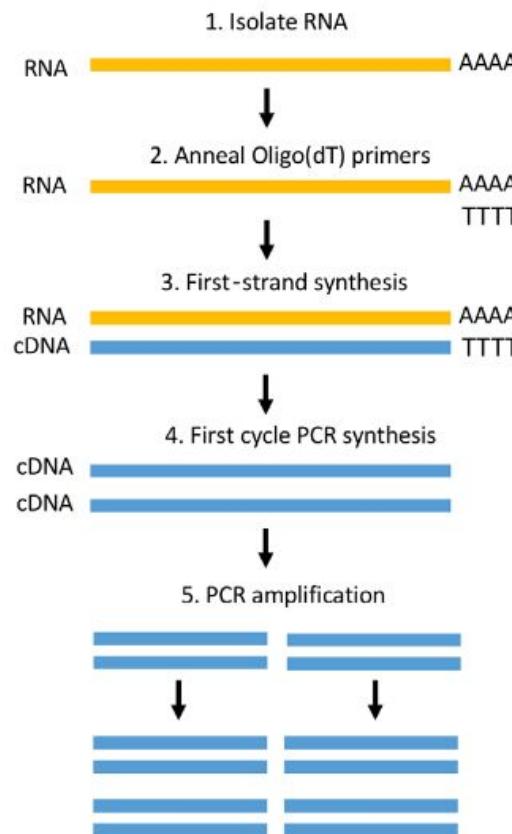
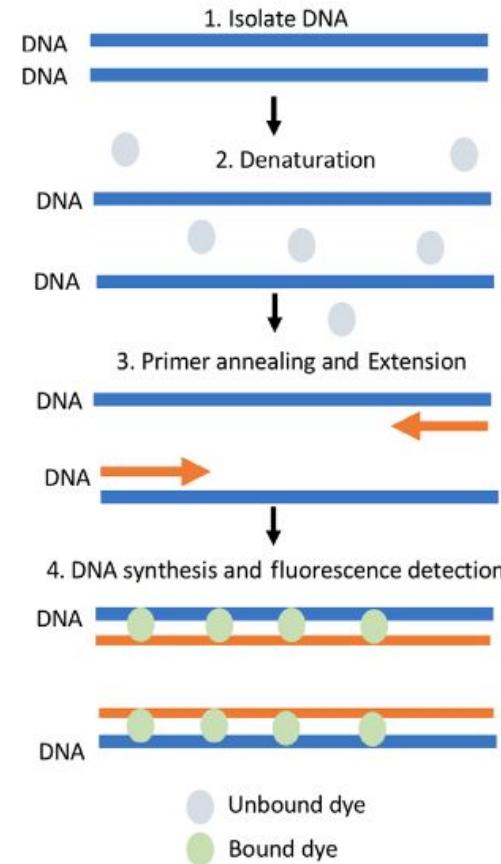
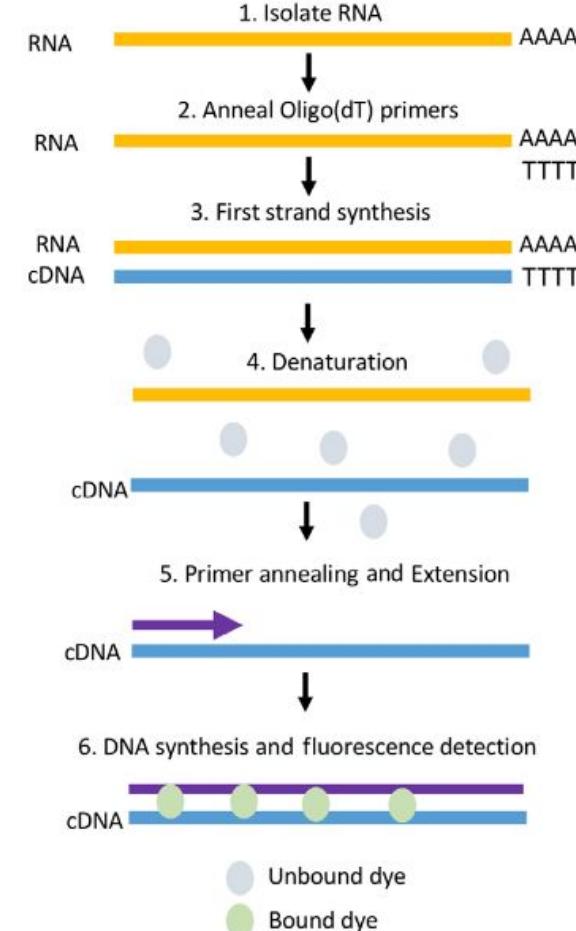


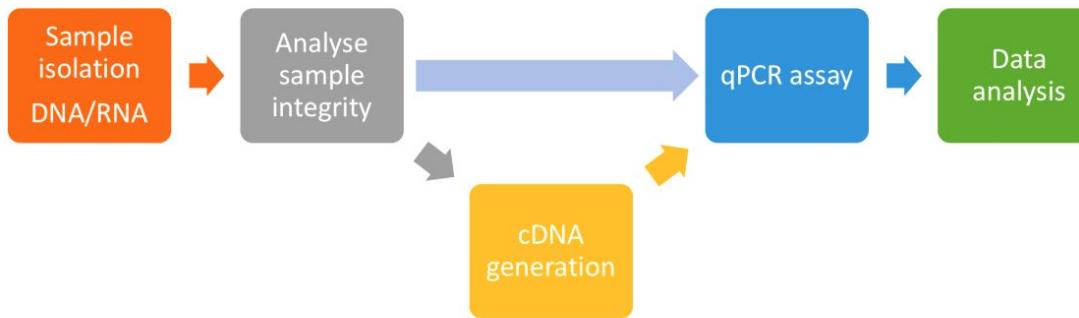
## Reverse transcription quantitative real-time PCR (RT-qPCR)



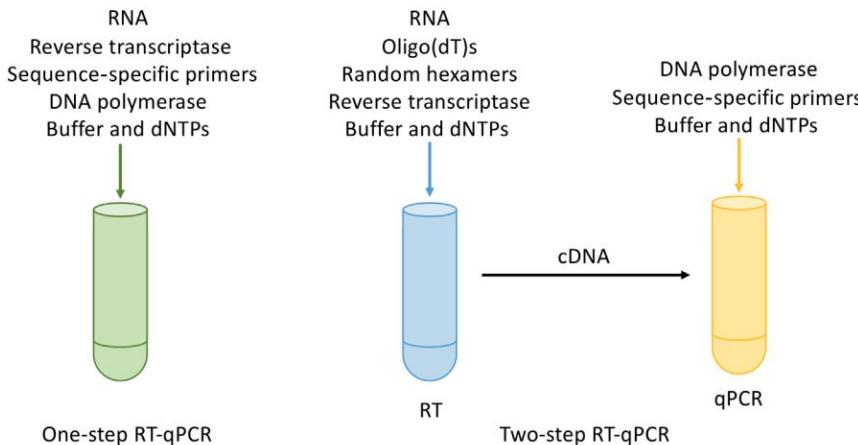
# Comparative measurement by RT-qPCR



**A****Reverse transcription-PCR****B****Quantitative real time – PCR (qPCR)****C****Reverse transcription quantitative real-time PCR (RT-qPCR)**

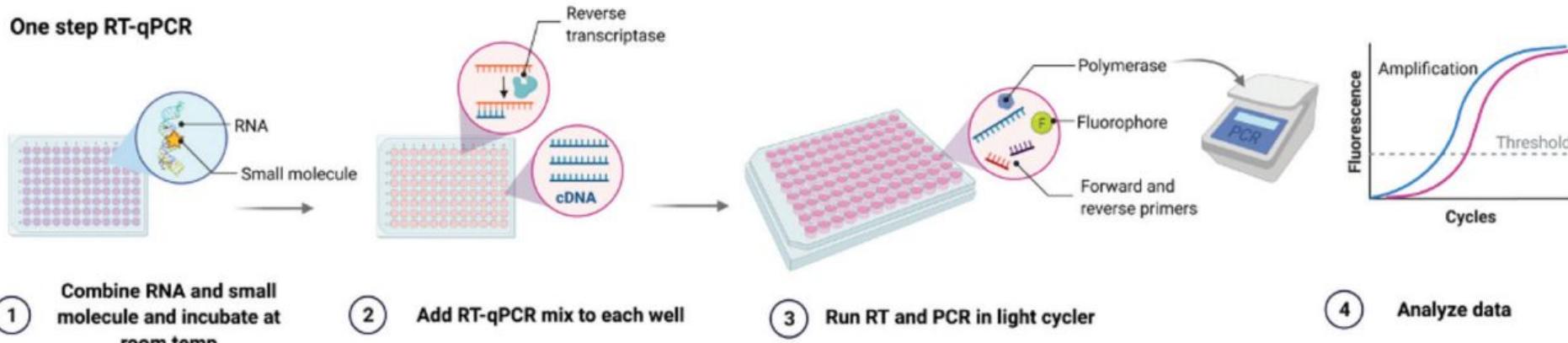


**Figure 2. Workflow of a standard qPCR and RT-qPCR experiment.** Following sample isolation, the integrity is analysed prior to cDNA generation and commencement of the qPCR assay using either intercalating dyes or hydrolysis probes. Fluorescence is detected throughout the PCR cycles and used to generate an amplification curve which is used to quantitate the target sample during data analysis.

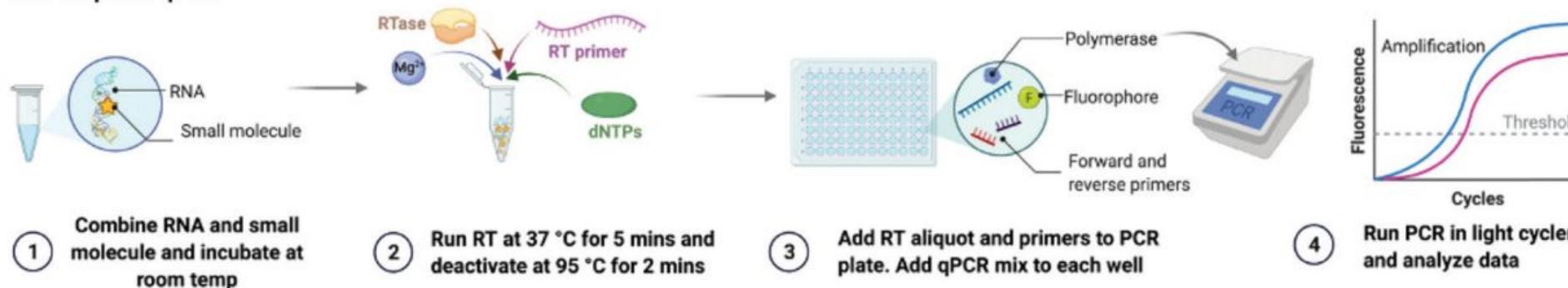


**Figure 3. One-step vs two-step RT-qPCR.** One-step RT-qPCR involves the generation of cDNA via reverse transcription and qPCR amplification of the target sequence in one reaction. Two-step RT-qPCR separates out the two steps (RT-PCR and qPCR), thus enabling more target sequences to be analysed in the qPCR reaction.

### One step RT-qPCR



### Two-steps RT-qPCR



# Measurements of Gene Expression

- RNA-based Measurements

- Low-Throughput
  - Northern Blot: Based on Hybridization technique and fluorescent probe
  - RT-PCR: Based on PCR and Reverse Transcriptase (RT)
  - RT-qPCR: Based on PCR with fluorescent dye
- High-Throughput
  - Microarray: Based on Hybridization technique and fluorescent probe
  - RNA-seq: Based on NGS

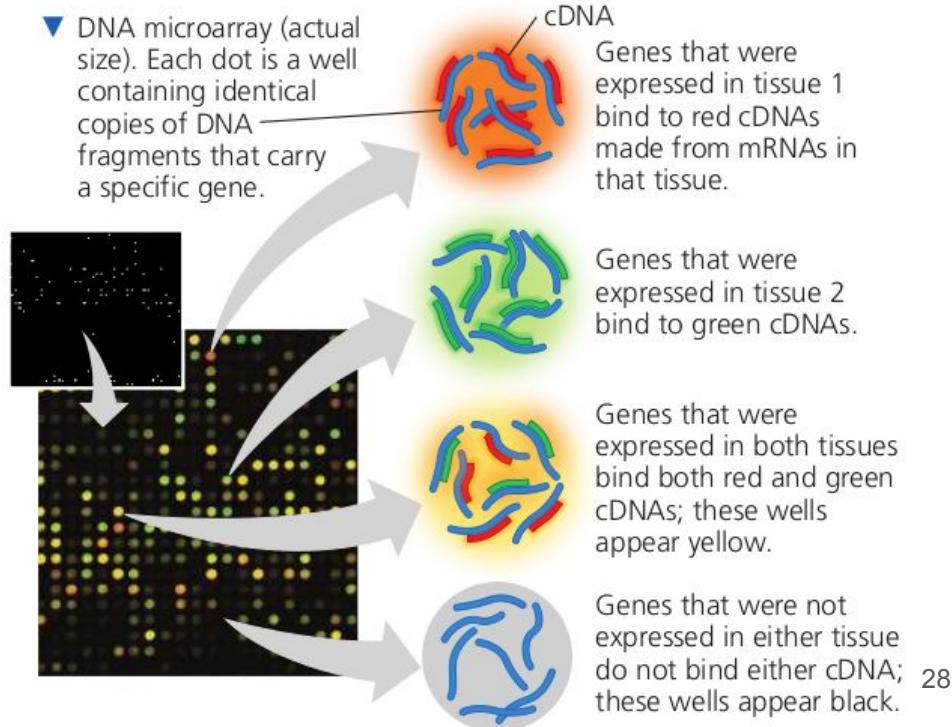
- Protein-based Measurements

- Low-Throughput
  - Western Blot: Based on antigen-antibody binding and PAGE
- High-Throughput
  - Mass Spectrometry
    - Two-Dimensional Polyacrylamide Gel Electrophoresis (2D-PAGE) -> MS
    - LC and HPLC -> MS

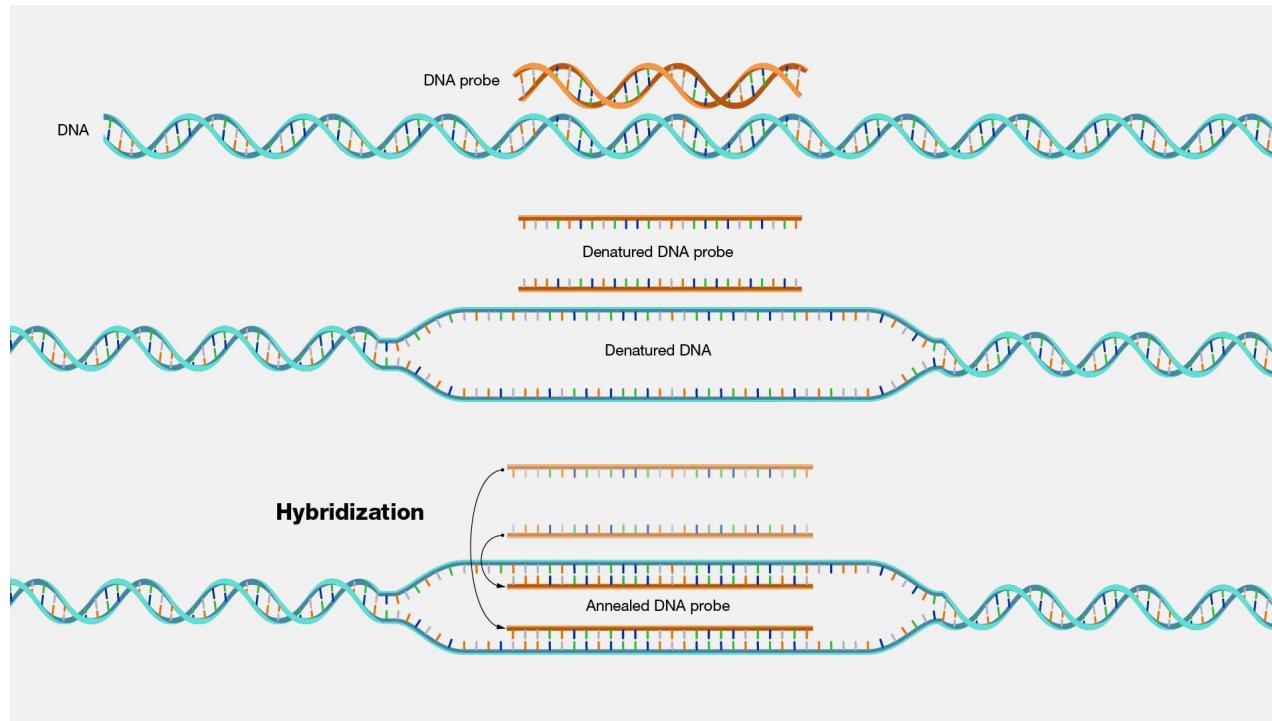
# Microarray Technology

- High-Throughput Technique for studying gene expression
- Based on Hybridization
- Have to know the DNA sequence of genes
- 2 systems in Microarray
  - Two color array
  - Single color array

▼ **Figure 20.13 Use of microarrays to analyze expression of many genes.** Researchers extracted mRNAs from two different human tissues and synthesized two sets of cDNAs, fluorescently labeled red (tissue 1) or green (tissue 2). Labeled cDNAs were hybridized with a microarray containing 5,760 human genes (about 25% of human genes), part of which is shown in the enlargement. Red indicates that the gene in that well was expressed in tissue 1, green in tissue 2, yellow in both, and black in neither. The fluorescence intensity at each spot indicates the relative expression of the gene.



# DNA Hybridization



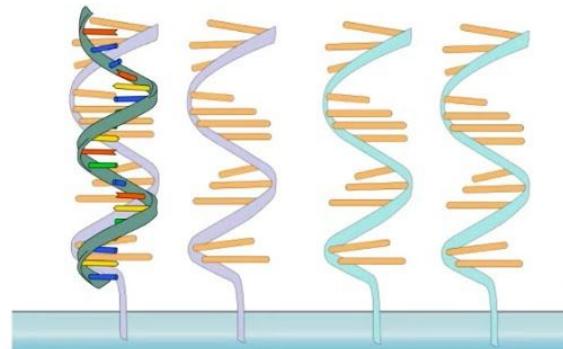
# Hybridization on Array

the cDNAs are labeled  
with **fluorescent tags**

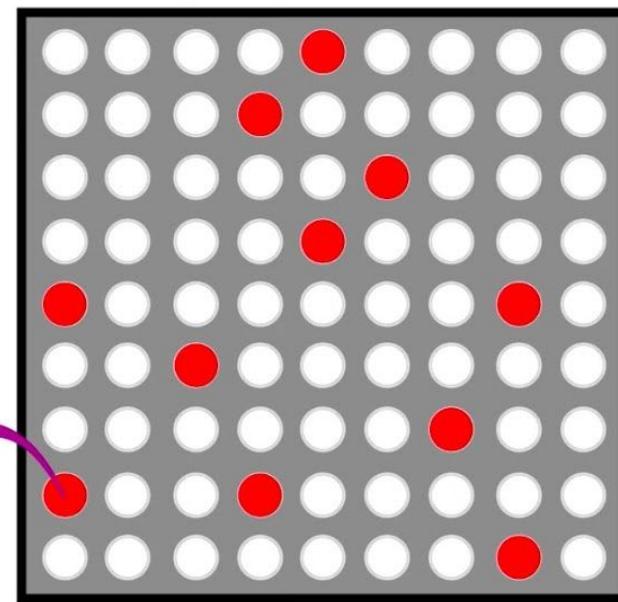


and introduced  
to the **array**

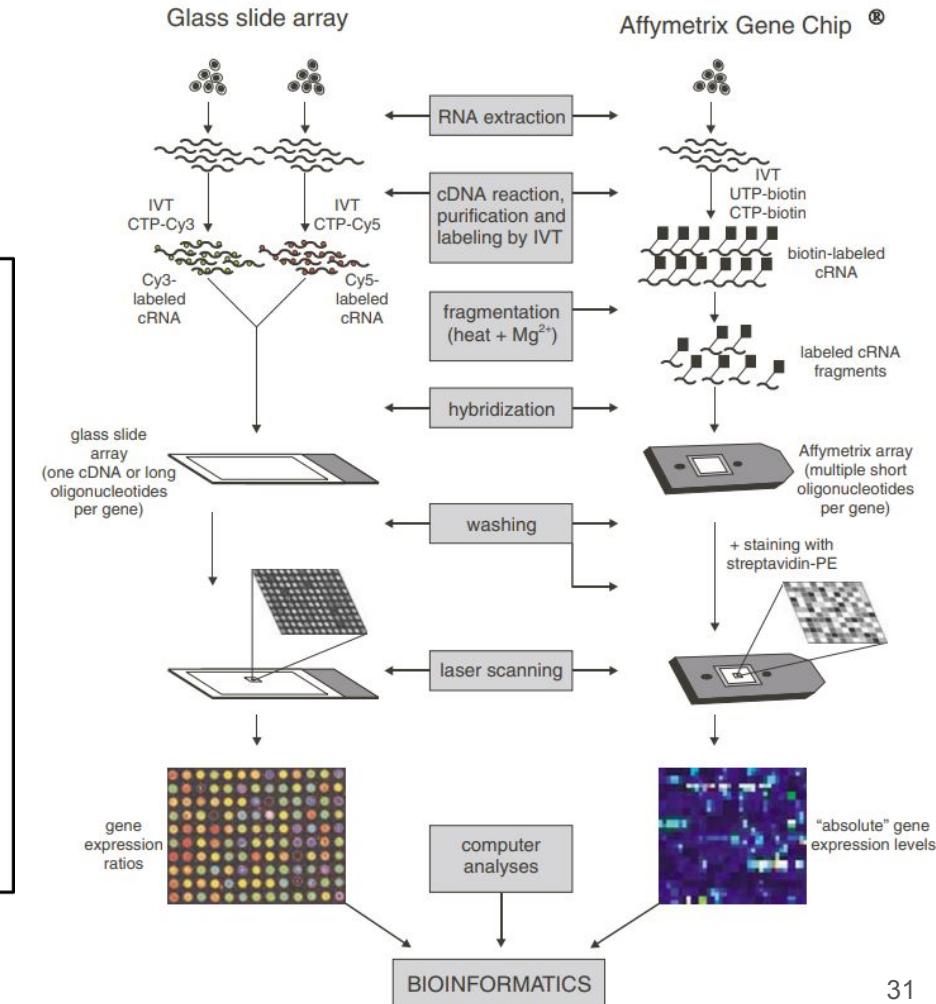
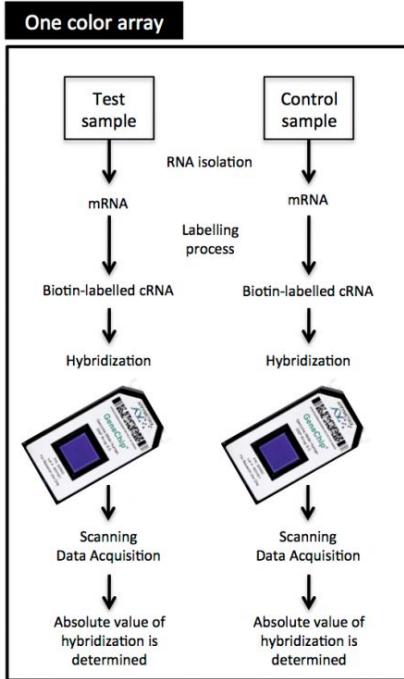
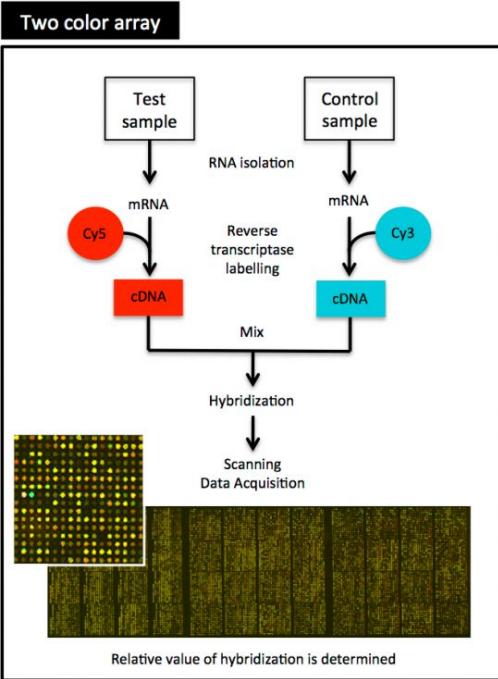
**fluorescence** is an  
indicator of binding



**binding means gene is expressed**

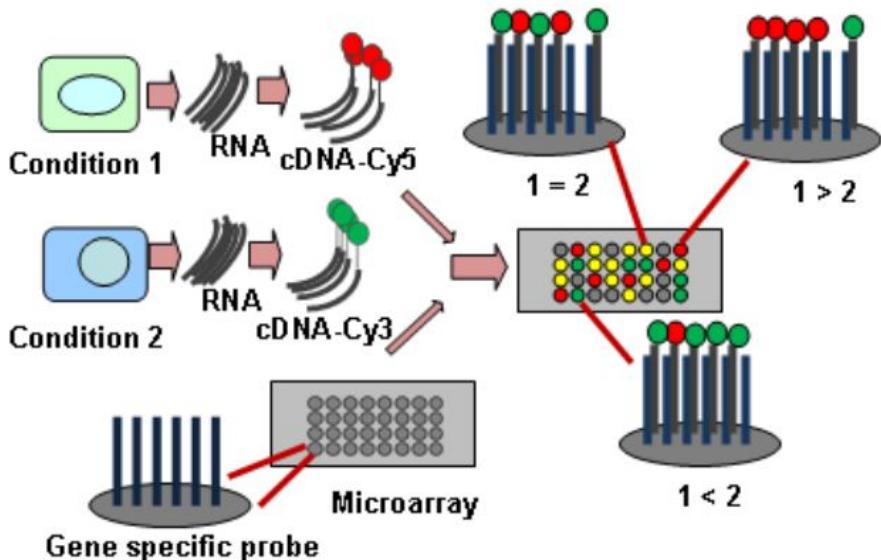


# 2 Systems of Microarray

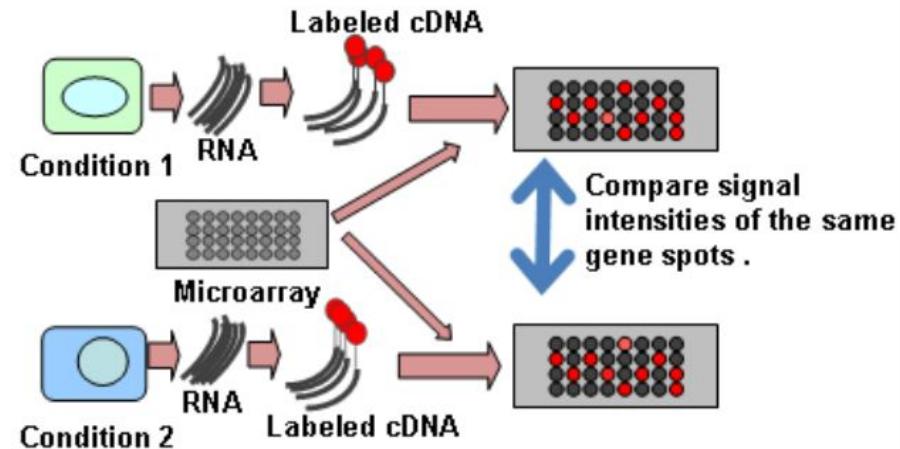


# 2 Systems of Microarray

Two Color Array

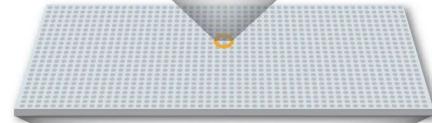
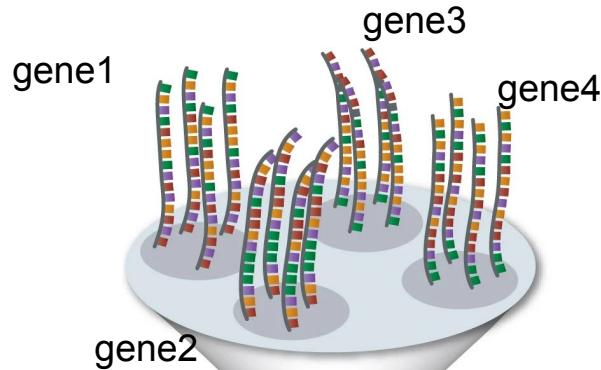


One Color Array



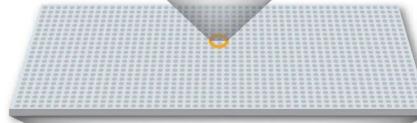
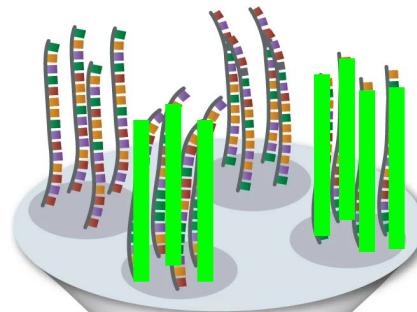
# Microarray: One Color

- One spot = One Gene



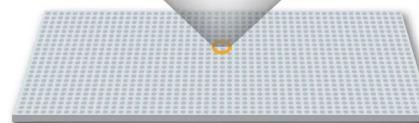
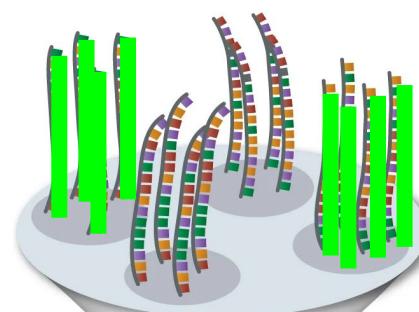
cDNA with label

Healthy



MICROARRAY

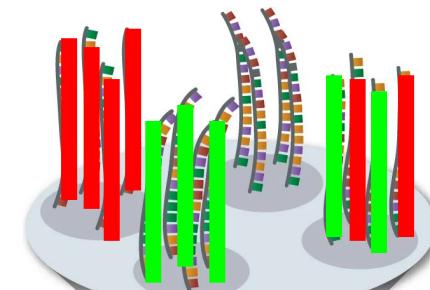
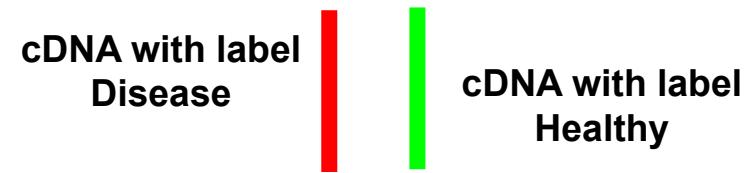
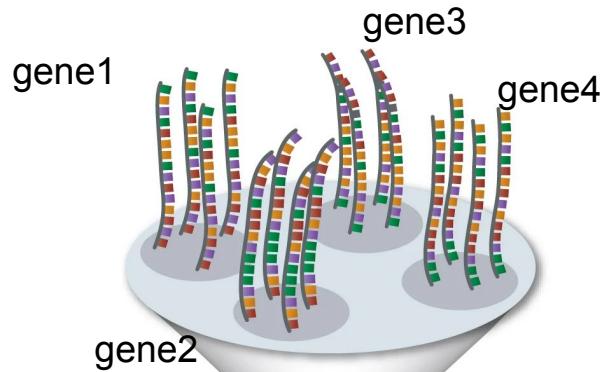
Disease



MICROARRAY

# Microarray: Two Color

- One spot = One Gene



# Measurements of Gene Expression

- RNA-based Measurements

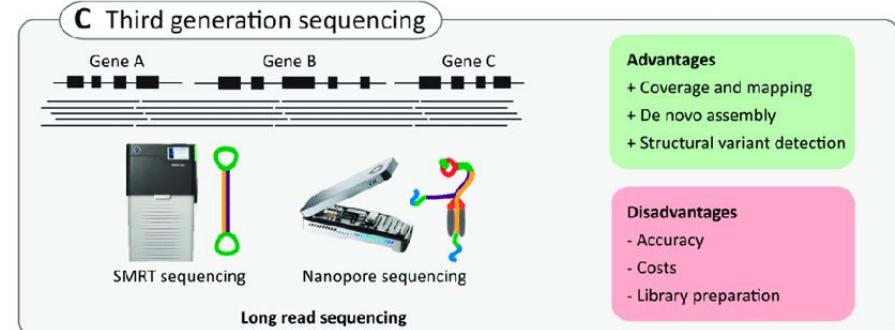
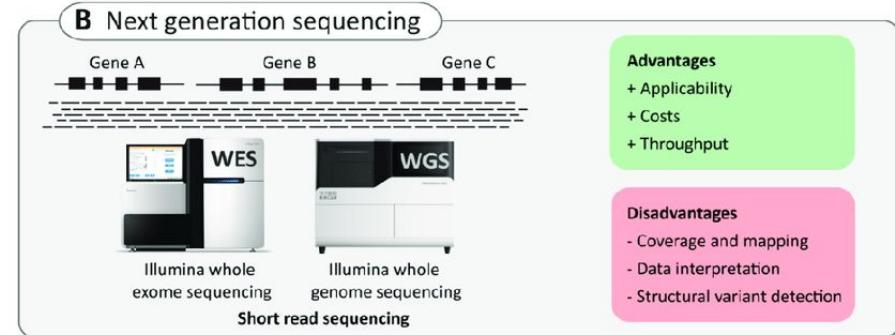
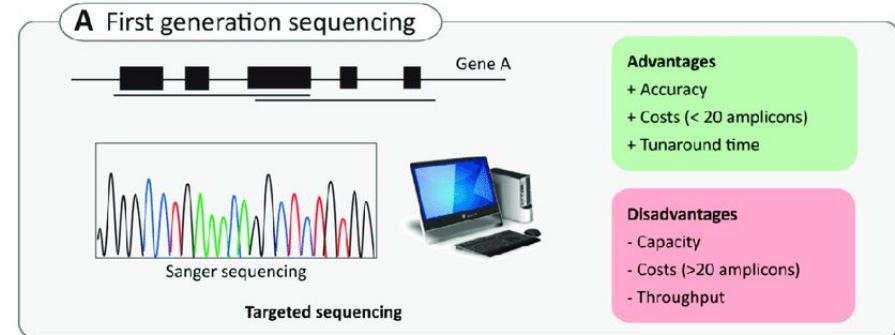
- Low-Throughput
  - Northern Blot: Based on Hybridization technique and fluorescent probe
  - RT-PCR: Based on PCR and Reverse Transcriptase (RT)
  - RT-qPCR: Based on PCR with fluorescent dye
- High-Throughput
  - Microarray: Based on Hybridization technique and fluorescent probe
  - RNA-seq: Based on NGS

- Protein-based Measurements

- Low-Throughput
  - Western Blot: Based on antigen-antibody binding and PAGE
- High-Throughput
  - Mass Spectrometry
    - Two-Dimensional Polyacrylamide Gel Electrophoresis (2D-PAGE) -> MS
    - LC and HPLC -> MS

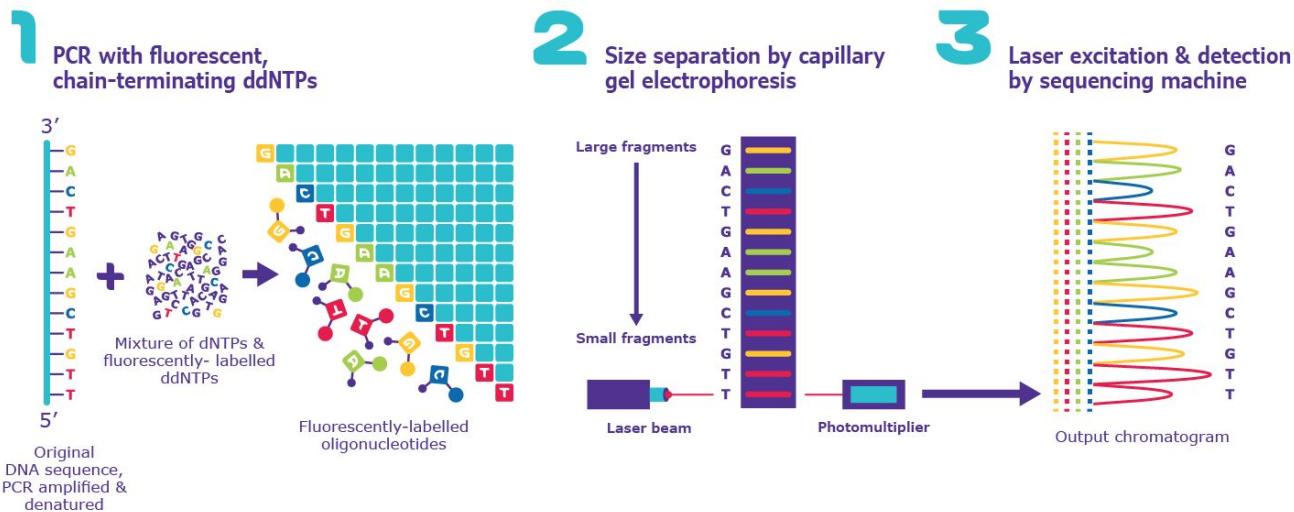
# Generations of DNA sequencing

- First Generation
  - Sanger sequencing
    - Targeted sequencing
    - 300-1000 bases
- Next generation sequencing (NGS)
  - Illumina
    - Short read (50-350 bases)
- Third generation sequencing
  - Oxford Nanopore sequencing
    - Long read (1–50 kb)
  - Pacific Bioscience (PacBio)
    - Long read (1–50 kb)



# Sequencing Technology

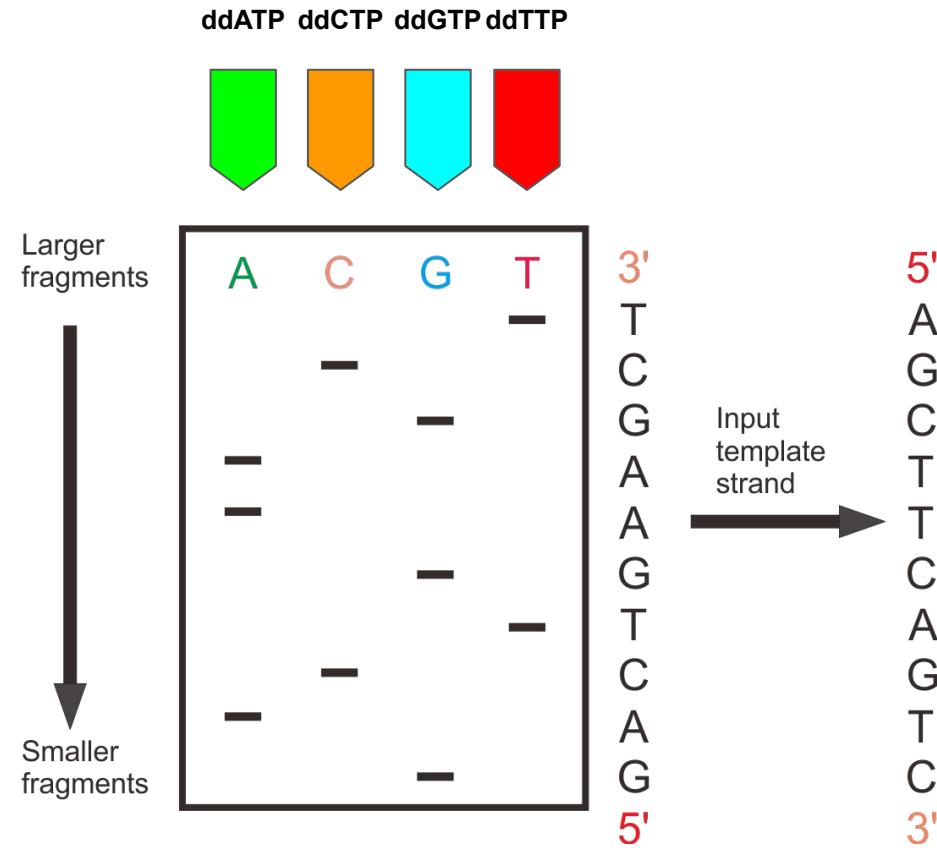
- Sanger Sequencing (Old school sequencing)
  - The Chain Termination Method
    - Dideoxyribonucleotides (ddNTPs)
    - Capillary Electrophoresis (size separation: short move fast; long move slow)



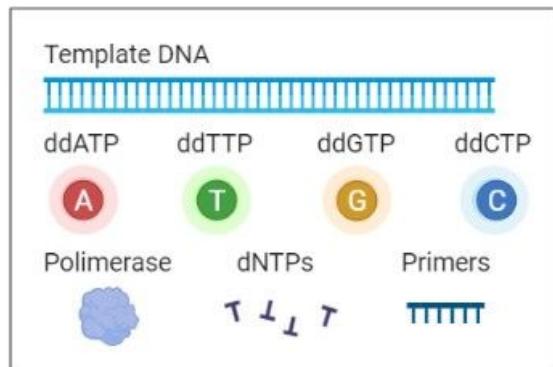
# Sanger Sequencing

- Conventional Sequencing

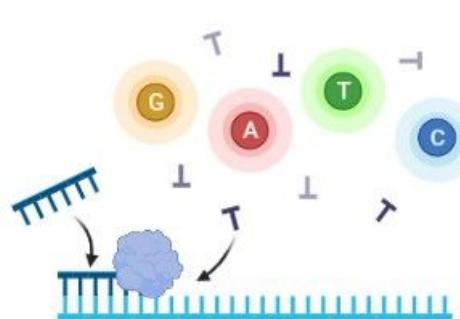
5' AGCTTCAGTC 3'  
G 5'  
AG 5'  
CAG 5'  
TCAG 5'  
GTCAG 5'  
AGTCAG 5'  
AAGTCAG 5'  
GAAGTCAG 5'  
CGAACAGTCAG 5'  
TCGAAGTCAG 5'



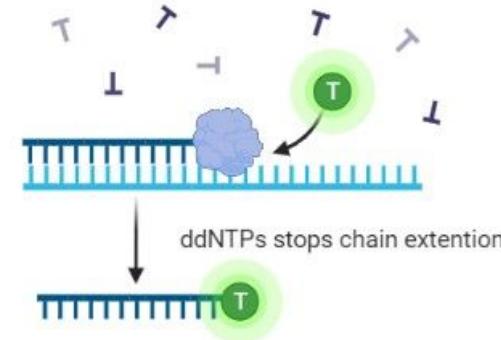
## Reagents



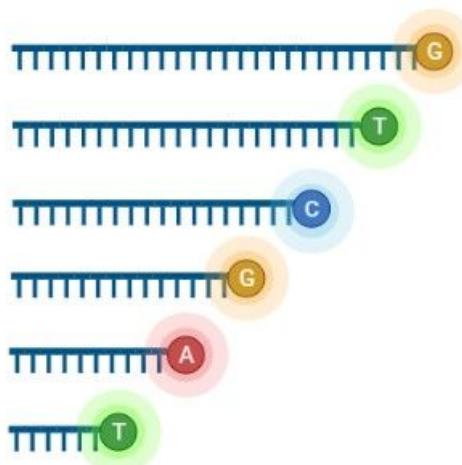
## ① Primer annealing and chain extension



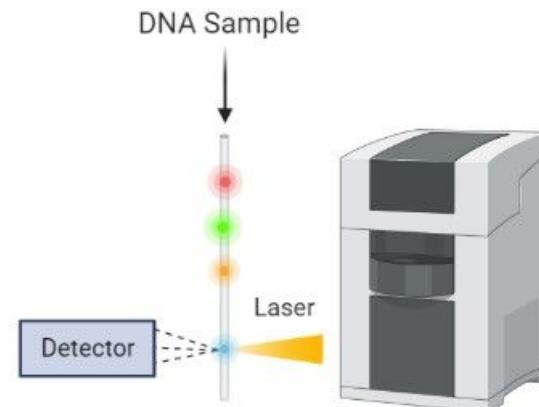
## ② ddNTP binding and chain termination



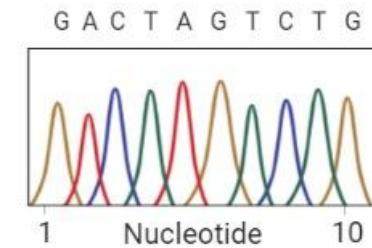
## ③ Fluorescently labelled DNA sample



## ④ Capillary gel electrophoresis and fluorescence detection



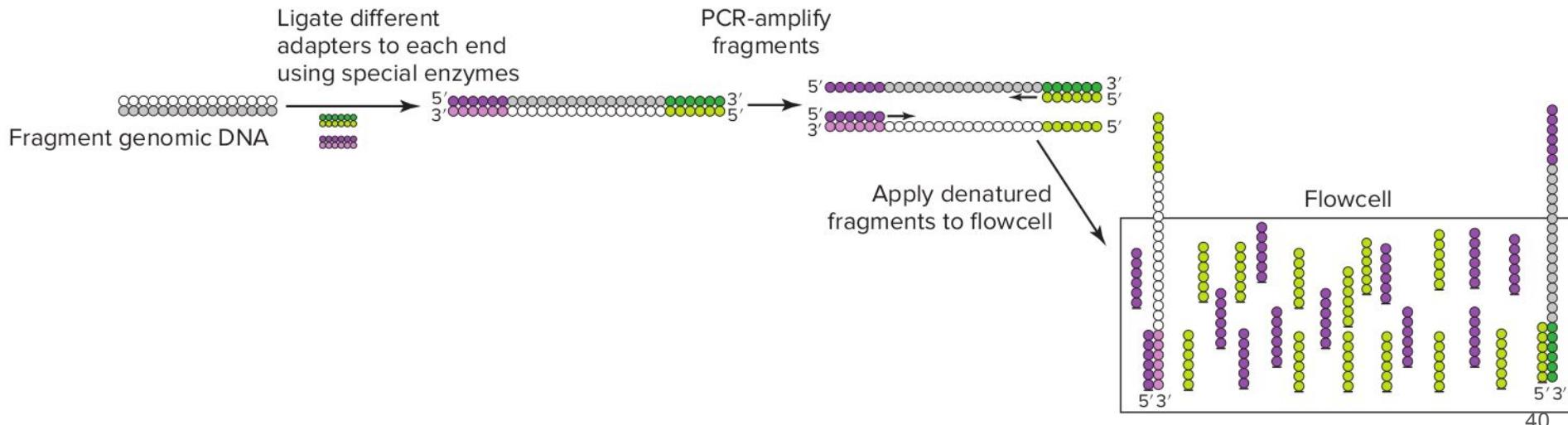
## ⑤ Sequence analysis and reconstruction



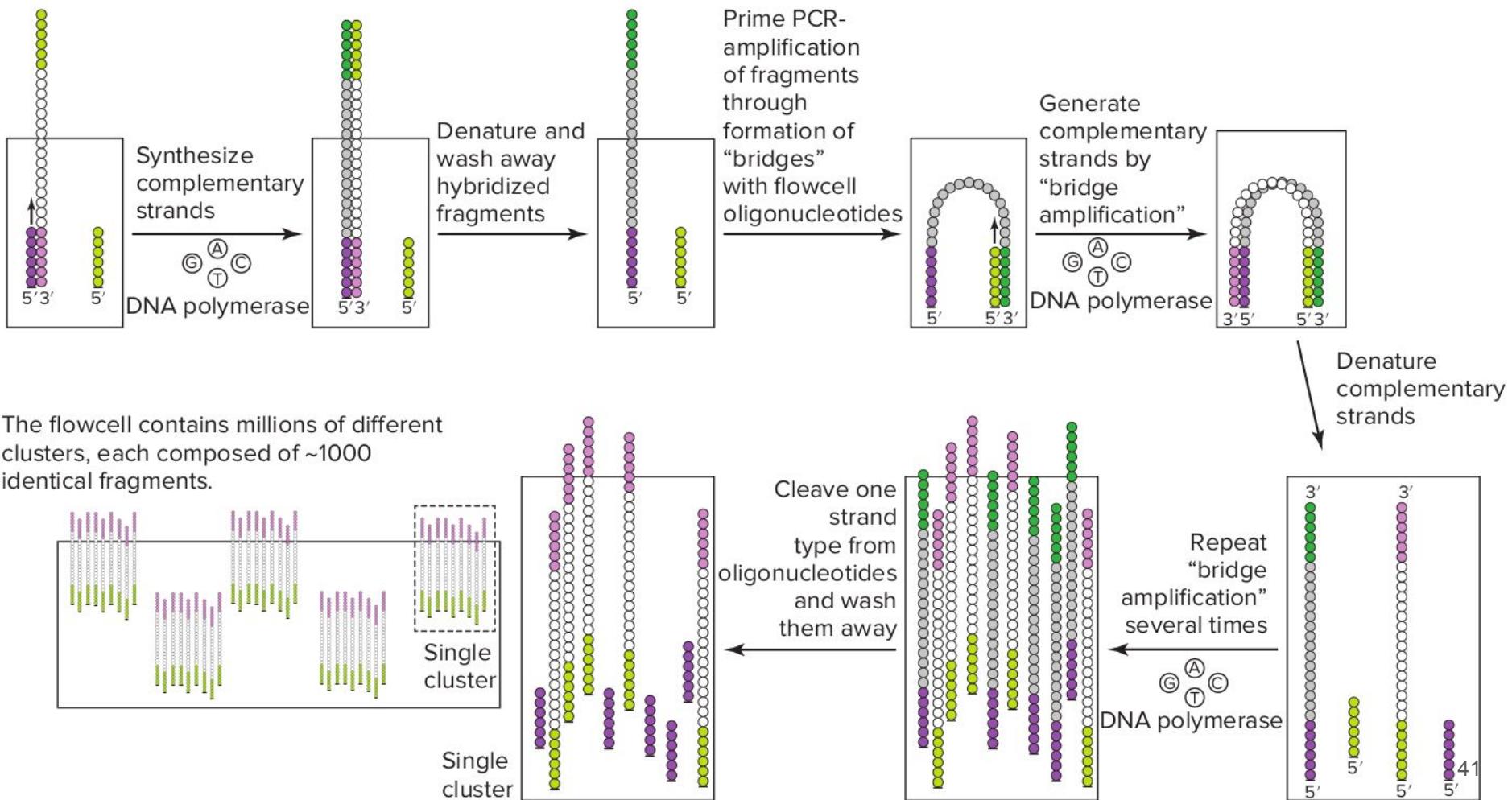
# Illumina sequencing

- Next Generation Sequencing (NGS)
- High-Throughput Sequencing
- Sequence by synthesis

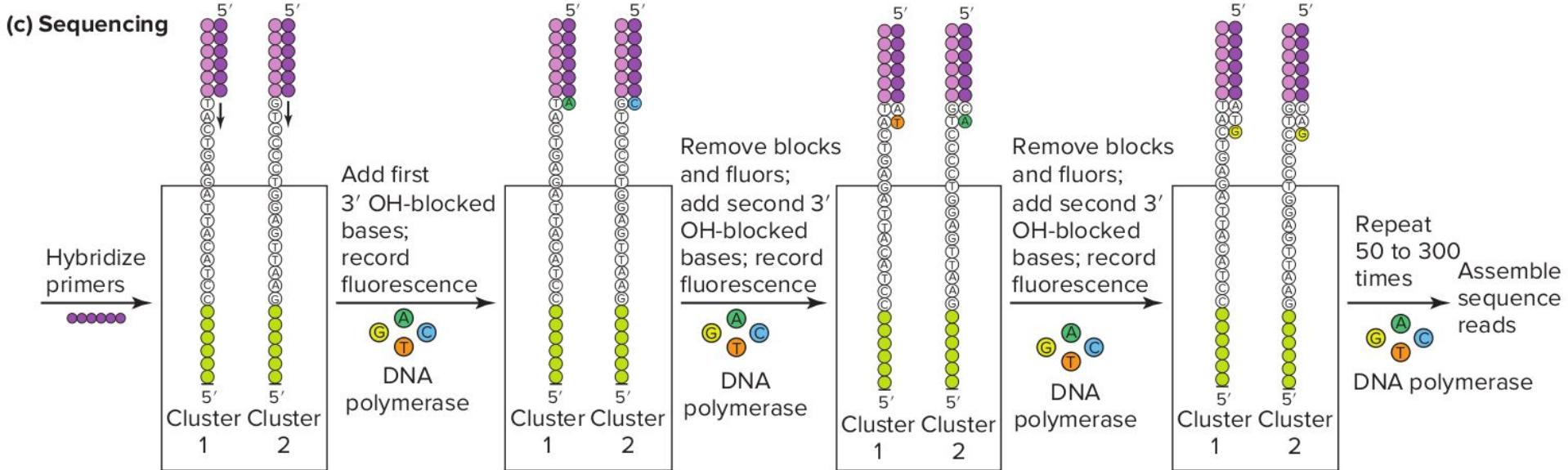
## (a) Sample preparation



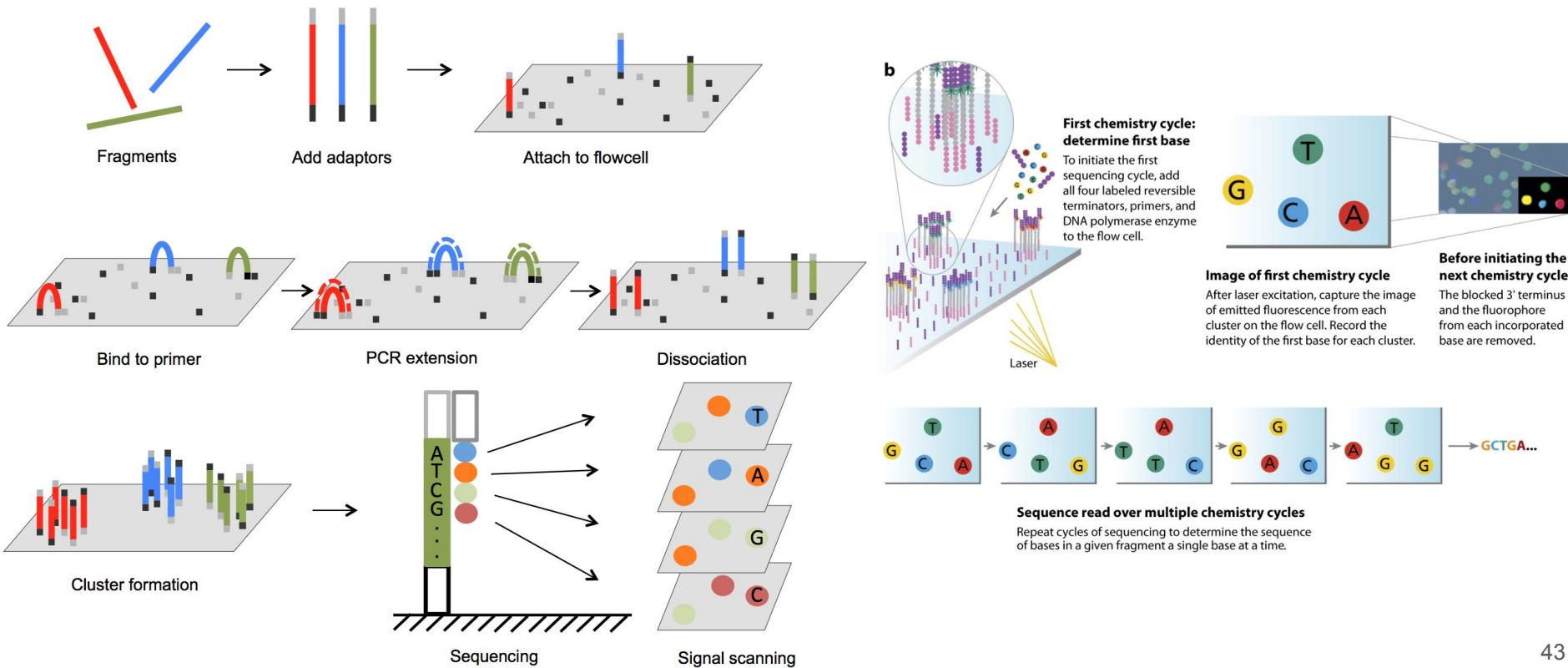
## (b) Cluster generation



# Illumina sequencing



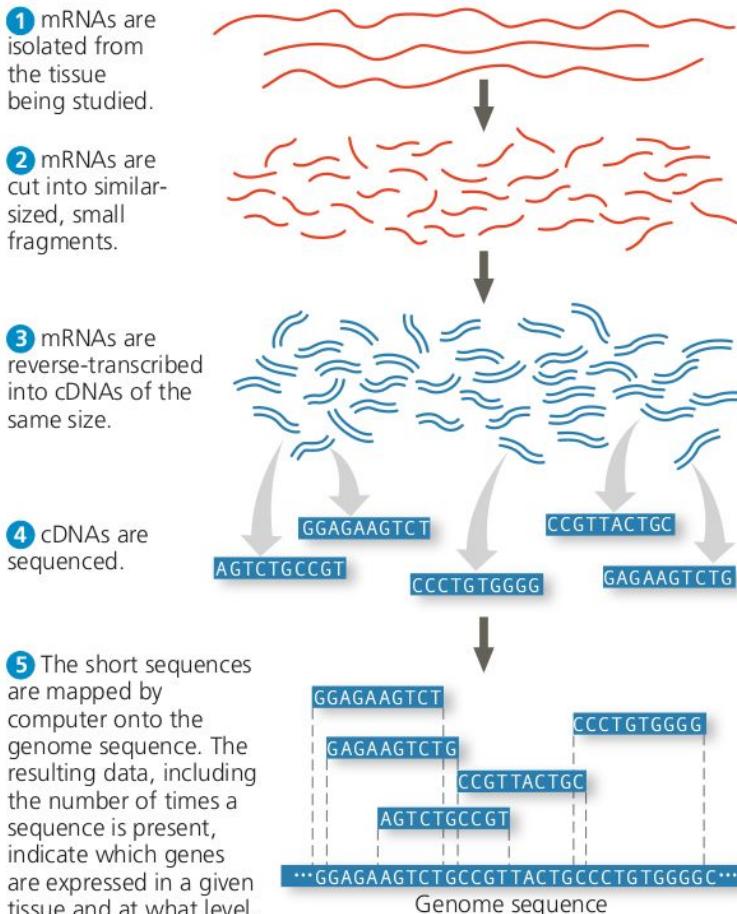
# Illumina sequencing



# RNA sequencing (RNA-seq)

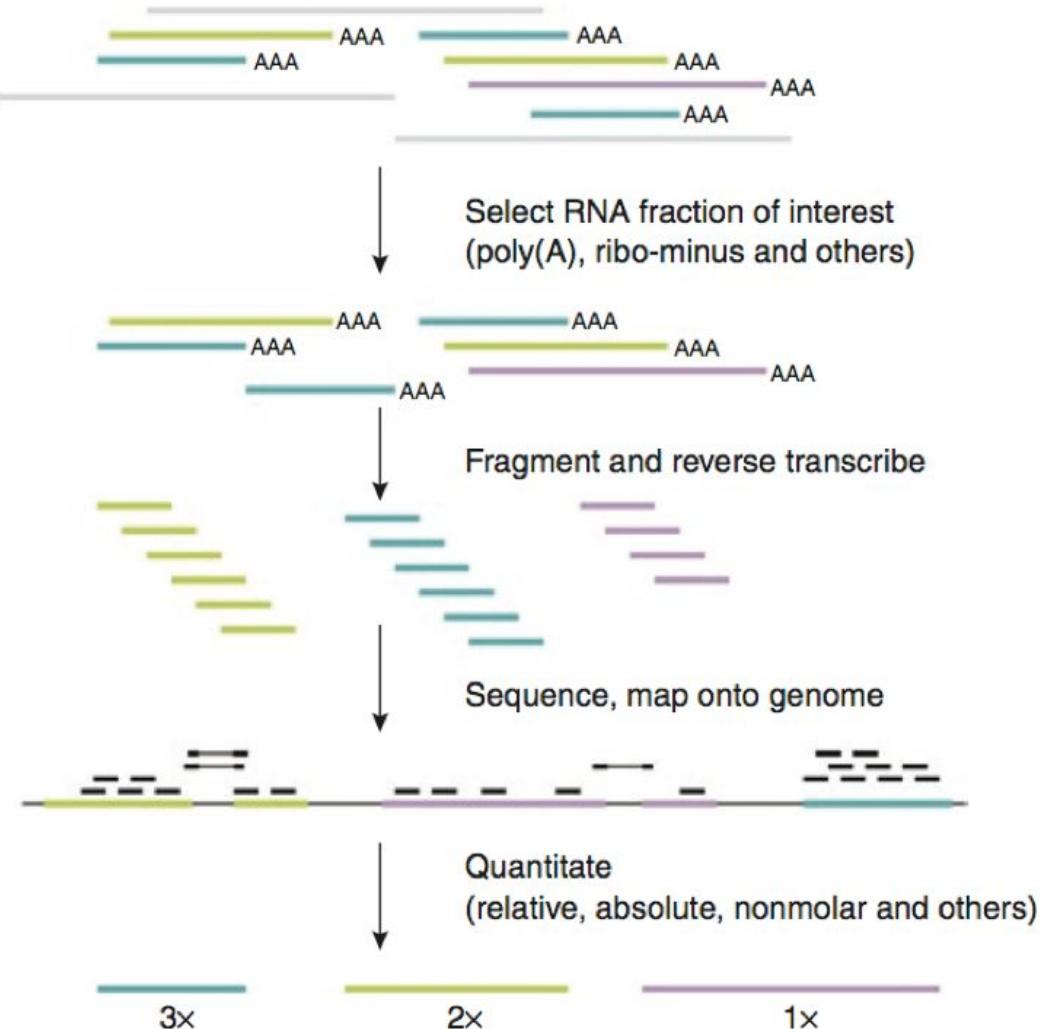
- Quantify all mRNA by Next Generation Sequencing (NGS)
- Can Detect all RNA in Cells (All genes)
- Required *in silico* analysis (Bioinformatics)
- Sequencing of all RNA in specific condition
- Use to study Transcriptome (all RNA: mRNA, rRNA, tRNA, and etc.)

▼ **Figure 20.12 Use of RNA sequencing (RNA-seq) to analyze expression of many genes.** RNA-seq yields a wide range of information about expression of genes, including their level of expression.

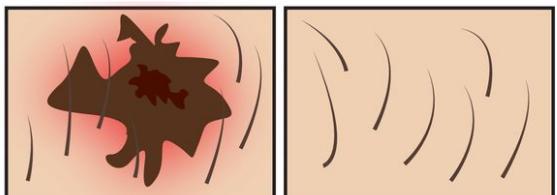


# RNA-seq

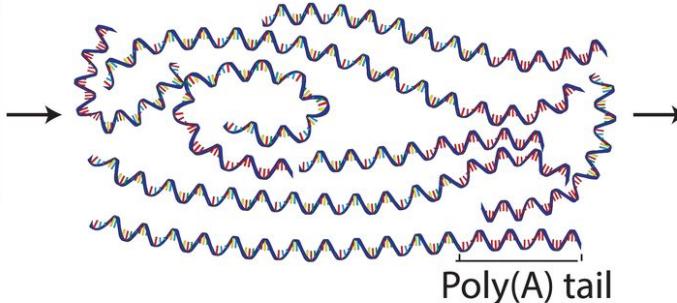
- RNA Preparation
- Sequencing
- *In silico*
  - Mapping
    - Alignment of each sequence to genome reference
  - Quantify sequence per gene
    - Count the number of sequences that aligned to each gene,



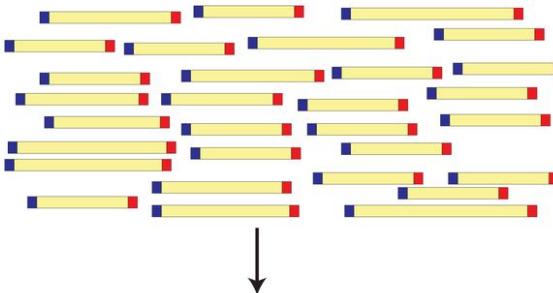
## Samples of interest



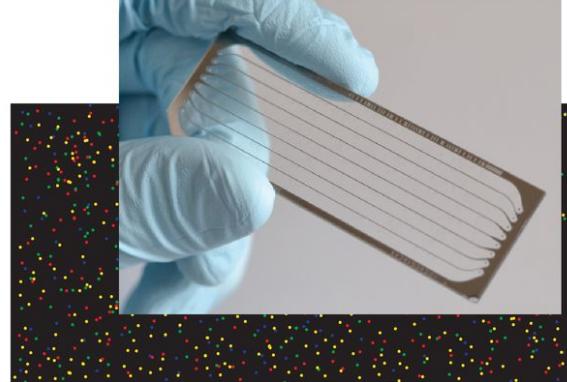
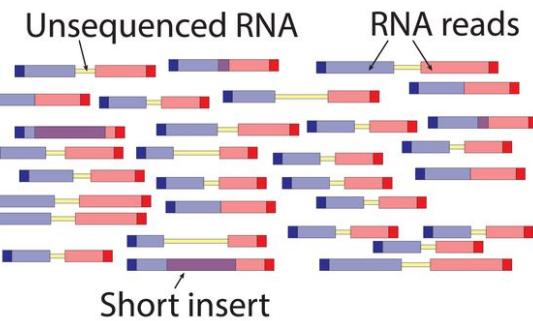
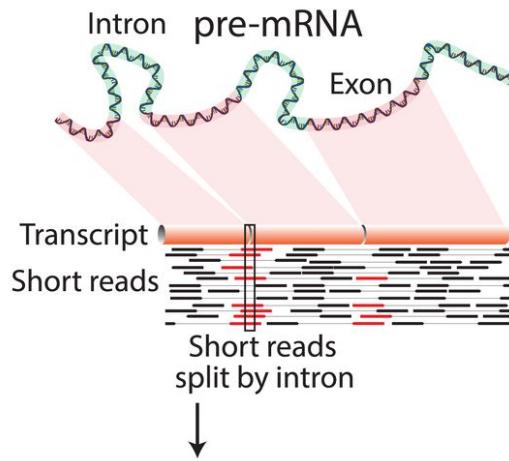
## Isolate RNAs



## Generate cDNA, fragment, size select, add linkers



Map to genome, transcriptome,  
and predicted exon junctions



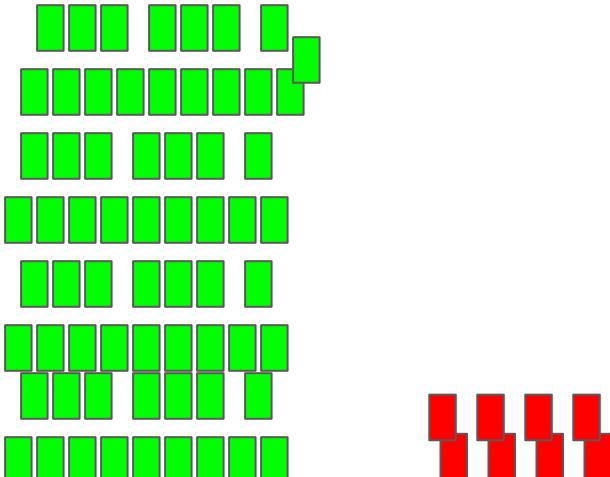
Downstream analysis

# RNA-seq

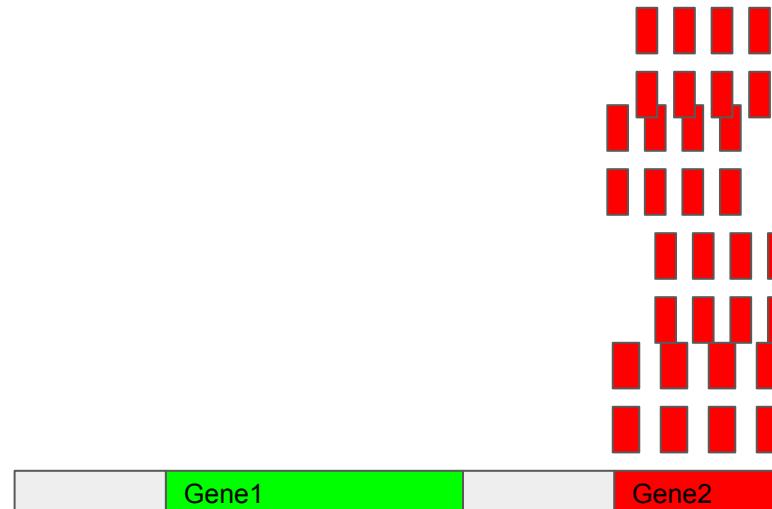


acggcatacagc tagtttcggaatca  
ttcggtccgga acggcatacagc

## Healthy

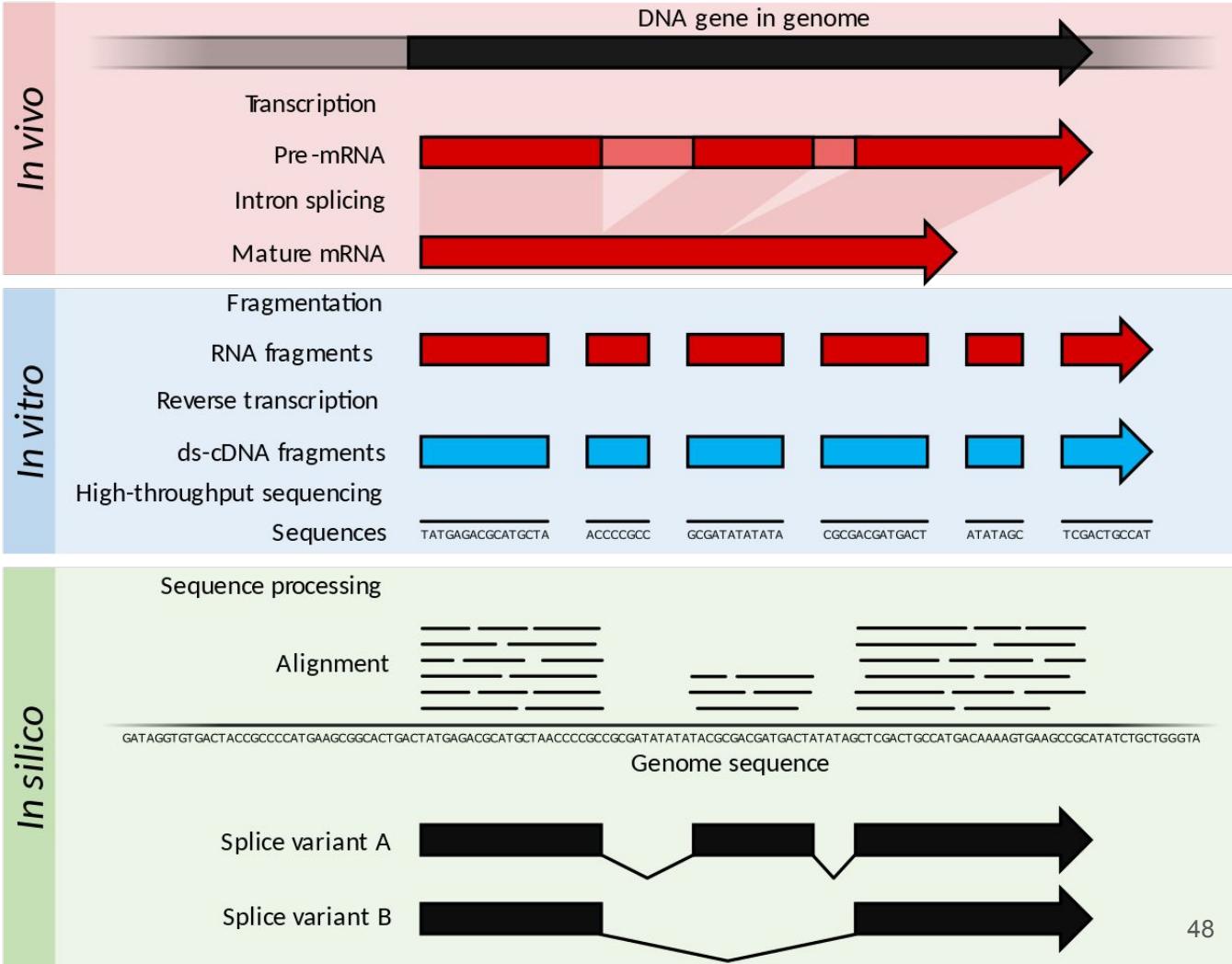


## Disease



# Summary of RNA-Seq

- *In vivo*
  - Perform in living thing
- *In vitro*
  - Perform in test tube
- *In silico*
  - Perform in computer (simulation)



# Measurements of Gene Expression

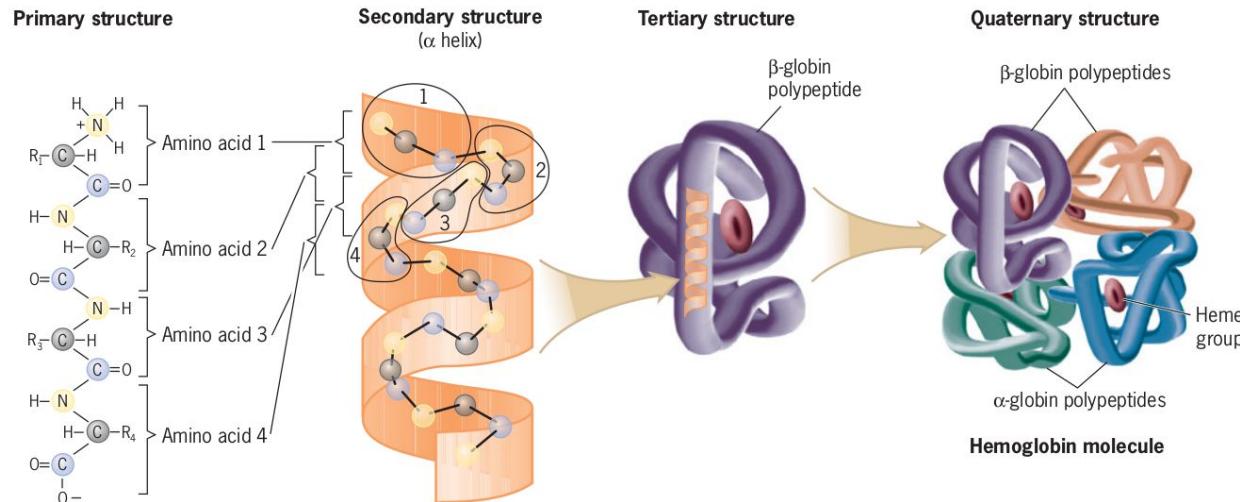
- RNA-based Measurements

- Low-Throughput
  - Northern Blot: Based on Hybridization technique and fluorescent probe
  - RT-PCR: Based on PCR and Reverse Transcriptase (RT)
  - RT-qPCR: Based on PCR with fluorescent dye
- High-Throughput
  - Microarray: Based on Hybridization technique and fluorescent probe
  - RNA-seq: Based on NGS

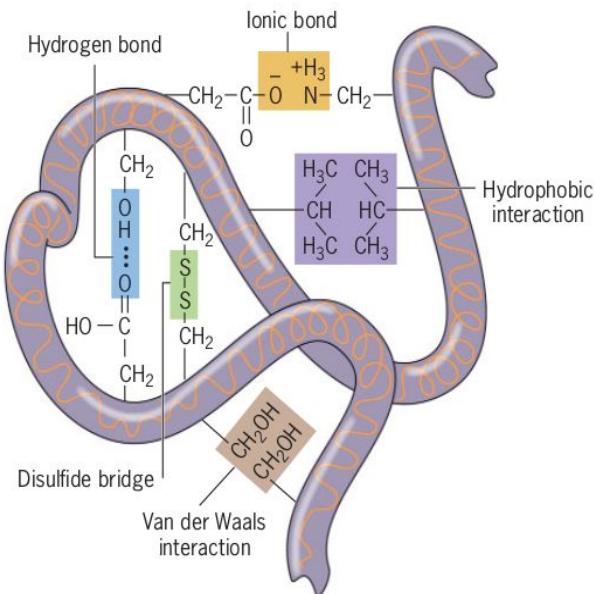
- Protein-based Measurements

- Low-Throughput
  - Western Blot: Based on antigen-antibody binding and PAGE
- High-Throughput
  - Mass Spectrometry
    - Two-Dimensional Polyacrylamide Gel Electrophoresis (2D-PAGE) -> MS
    - LC and HPLC -> MS

# Polypeptide and Protein Structure



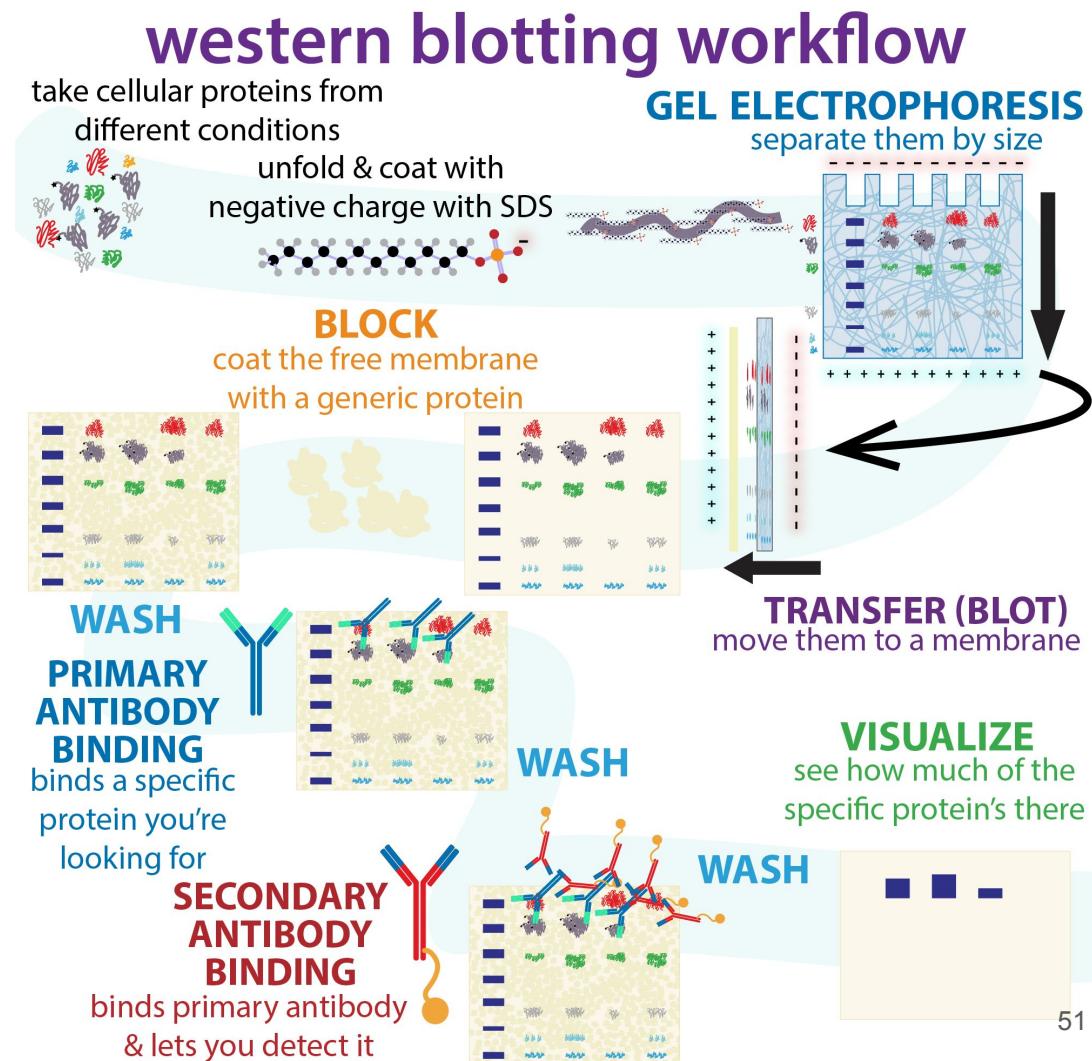
■ FIGURE 12.3 The four levels of organization in proteins—(1) primary, (2) secondary, (3) tertiary, and (4) quaternary structures—are illustrated using human hemoglobin as an example.



■ FIGURE 12.4 The five types of molecular interactions that determine the tertiary structure, or three-dimensional conformation, of a polypeptide. The disulfide bridge is a covalent bond; all other interactions are noncovalent.

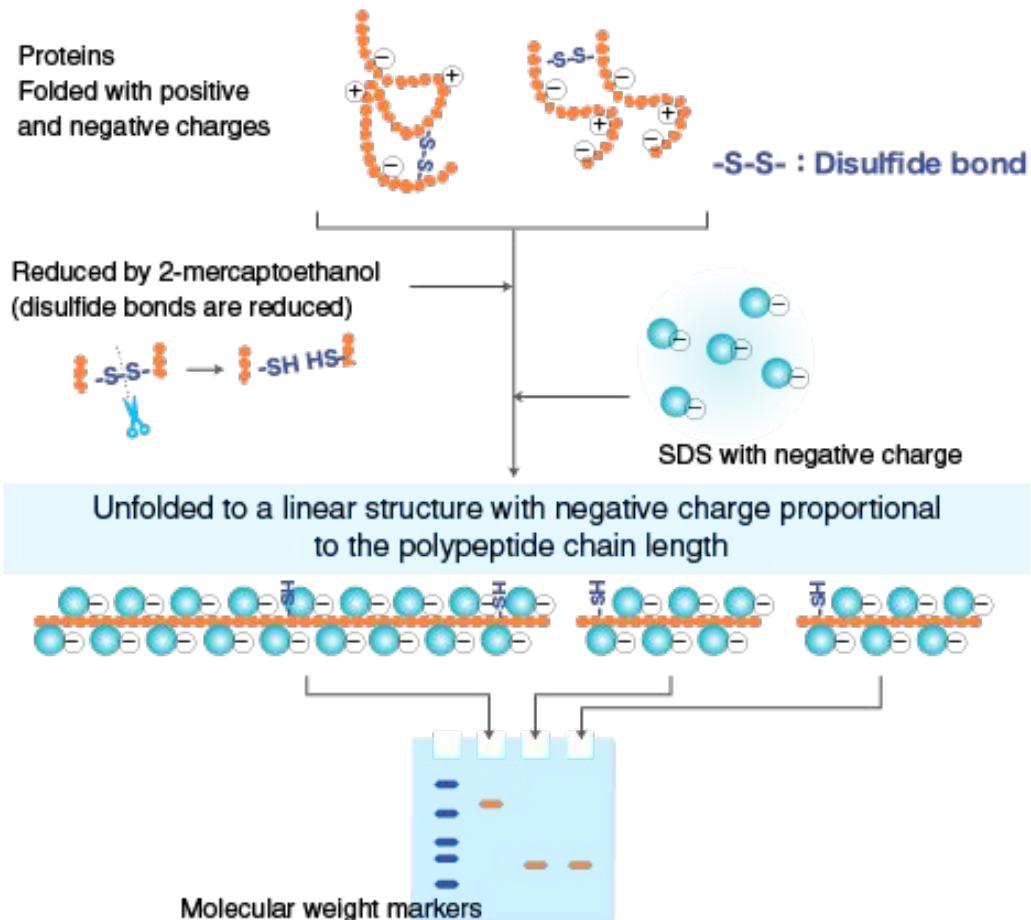
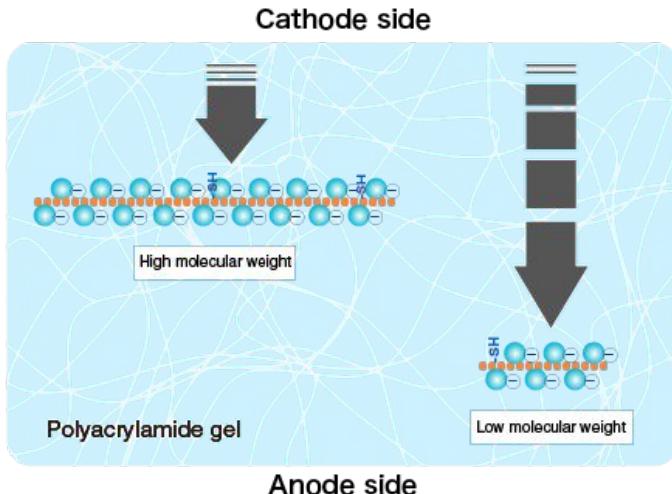
# Protein analysis by Western Blot

- Protein Extraction
- Protein Separation by SDS-PAGE
- Blotting: Transfer separated protein from gel to membrane
- Blocking
- Probe by Antibody
  - Primary
  - Secondary
- Visualization



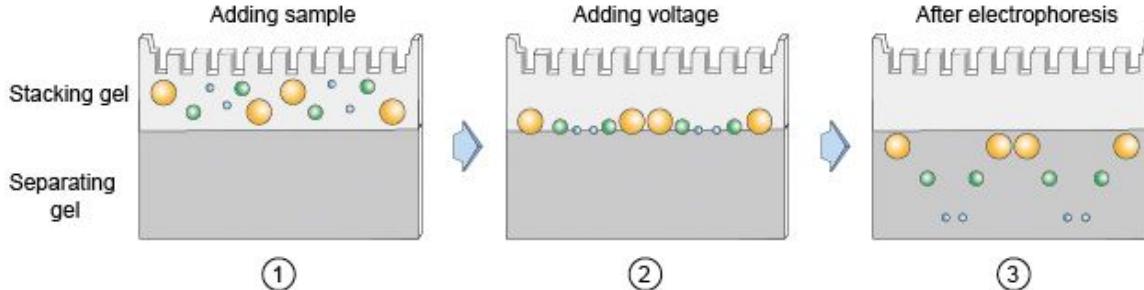
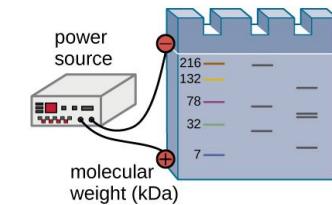
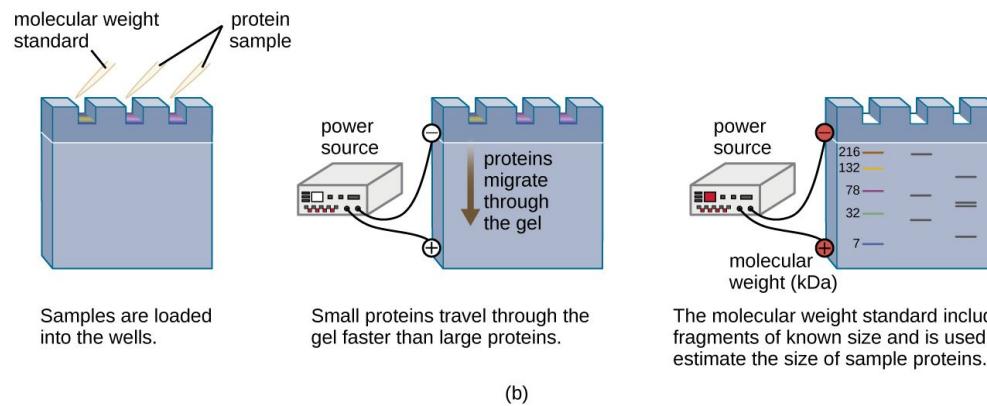
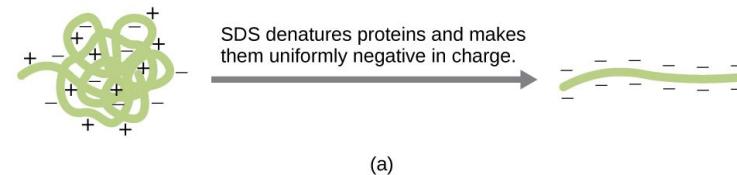
# SDS-PAGE

- Sodium Dodecyl Sulfate-PolyAcrylamide Gel Electrophoresis (**SDS-PAGE**)
- Separate Protein by Molecular Weight



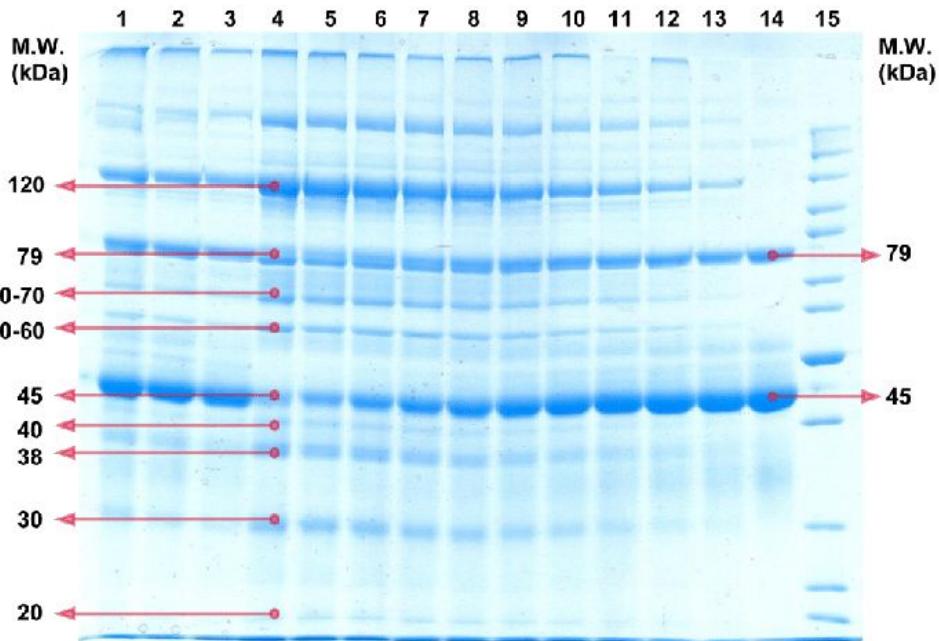
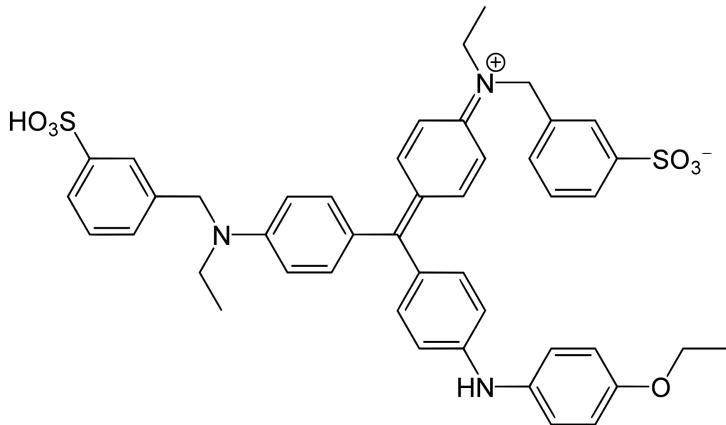
# SDS-PAGE

- Denature Protein to Primary Structure by Heat and Beta-Mercaptoethanol
- Make Negative Charge of protein by SDS
- Size separation by PAGE
  - Stacking gel
  - Resolving gel



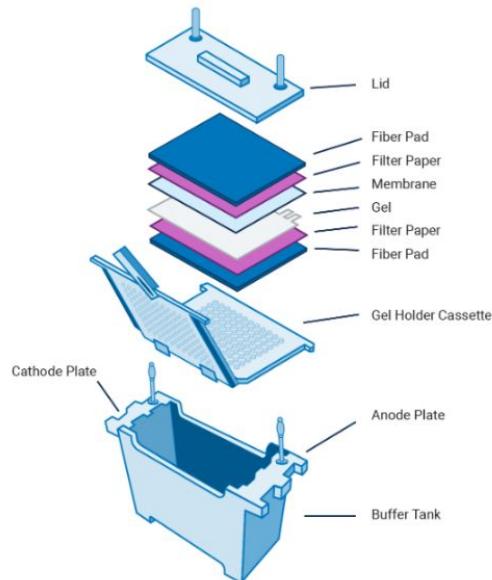
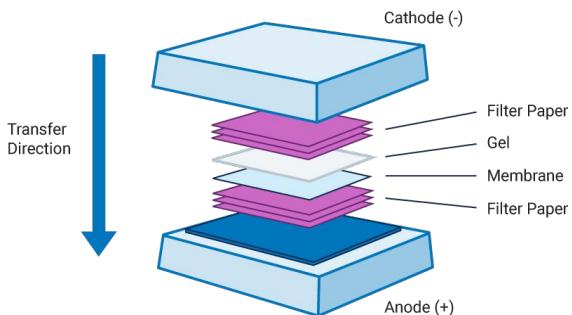
# SDS-PAGE

- Coomassie blue staining
  - Coomassie brilliant blue



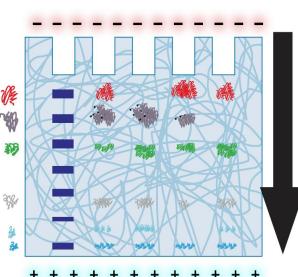
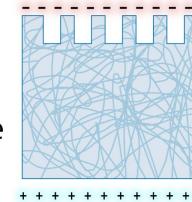
# Blotting

- Transfer Protein from gel to membrane
  - Semi-dry Transfer
  - Wet-tank Transfer



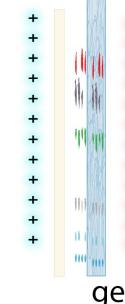
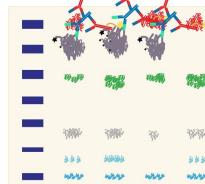
once you give the proteins the negative coat you can direct them to go towards a positive charge

First send proteins vertically through the gel



membrane

then use antibodies to probe the membrane to see how much of a SPECIFIC protein's there



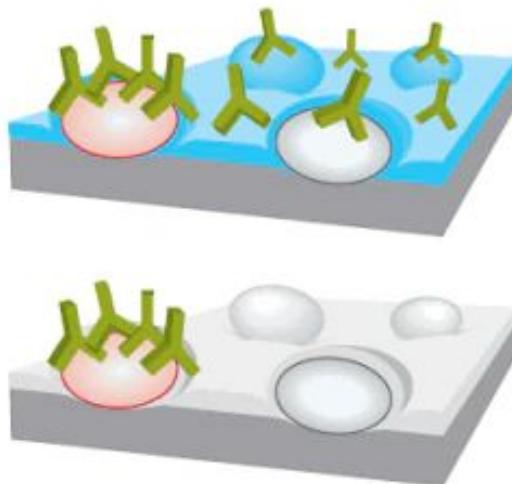
gel

then transfer them horizontally to a membrane if you used a pre-stained ladder you should see it

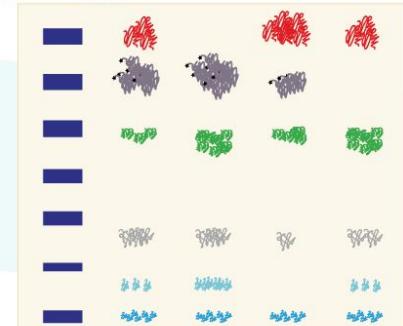
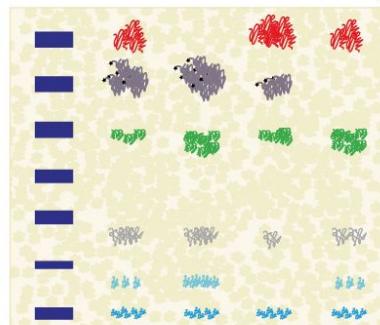
but you won't see the proteins in your sample

# Blocking

- Enhance specific binding of **Antibody** to our interesting protein
- Reduce noise from unwanted binding of antibody to other proteins

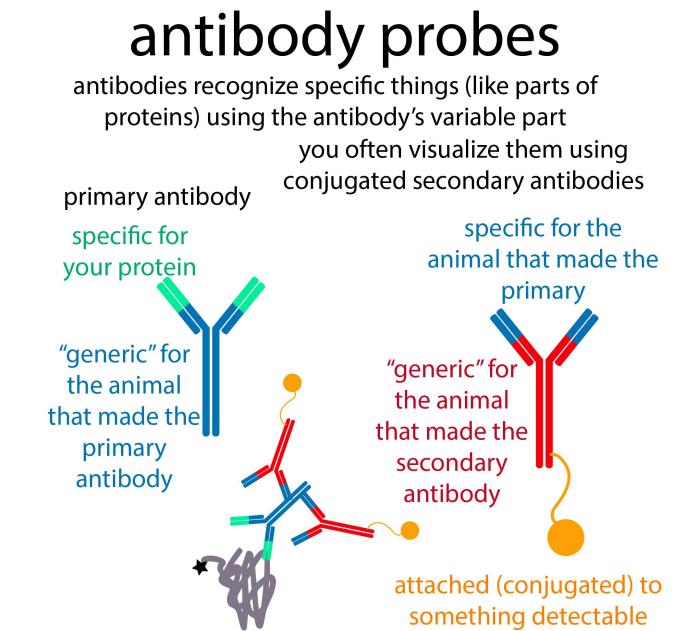
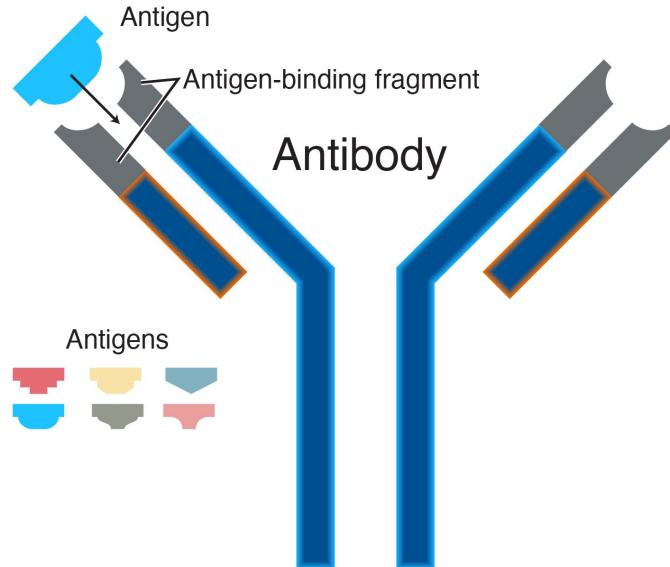


**BLOCK**  
coat the free membrane  
with a generic protein

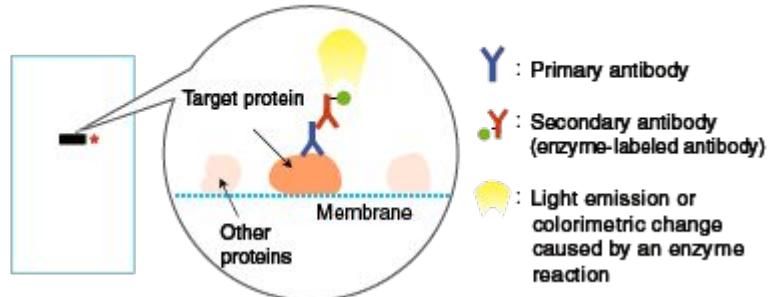


# Protein Detection

- Use the idea of **antigen** and **antibody** to Detect specific Protein

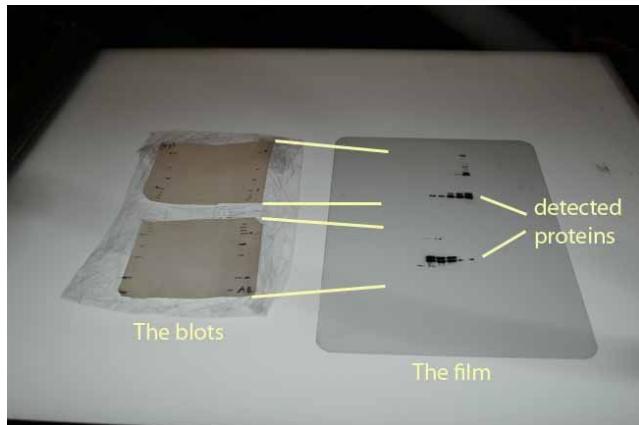
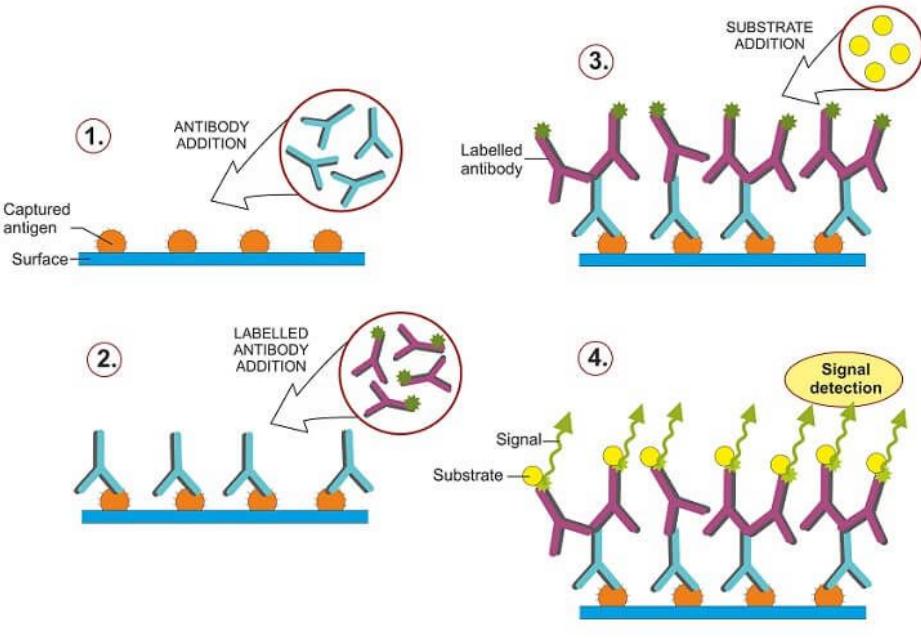


Probing with antibodies, and detection of the target protein by an enzyme reaction.



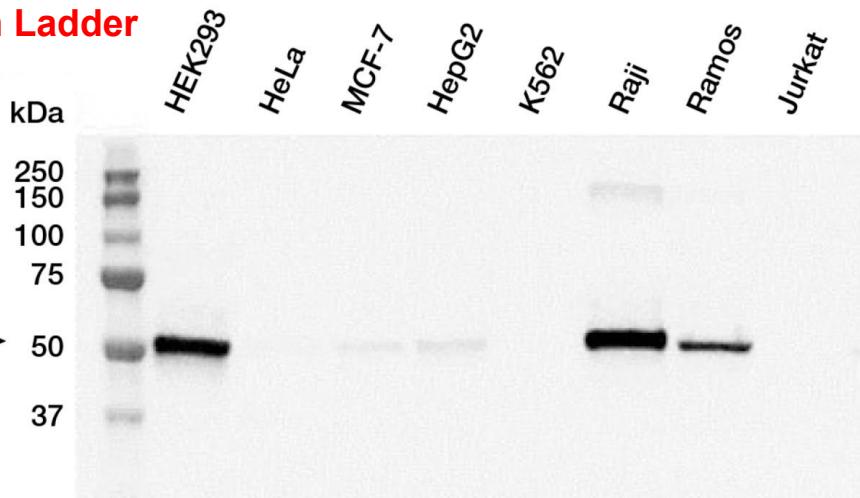
# Probe and Visualization

- Chemiluminescence Method
- Start Reaction by add **Substrate**
- The enzyme that attached to secondary antibody will react with substrate and produce the **light** or etc.
- Use **Film** or **Camera** to detect **light emission** as the product from chemical reaction between enzyme and substrate.

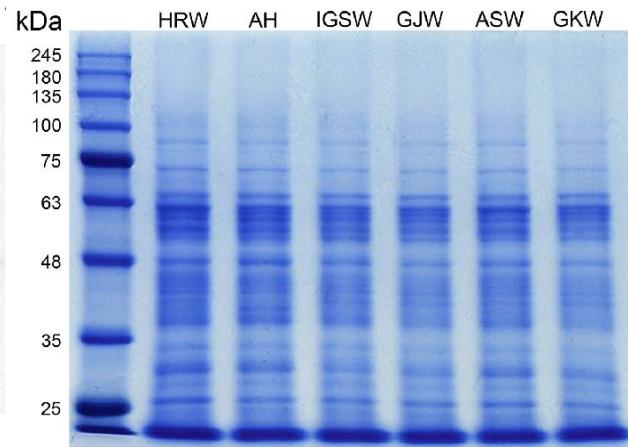


# Western blot Result

Protein Ladder



Coomassie blue staining



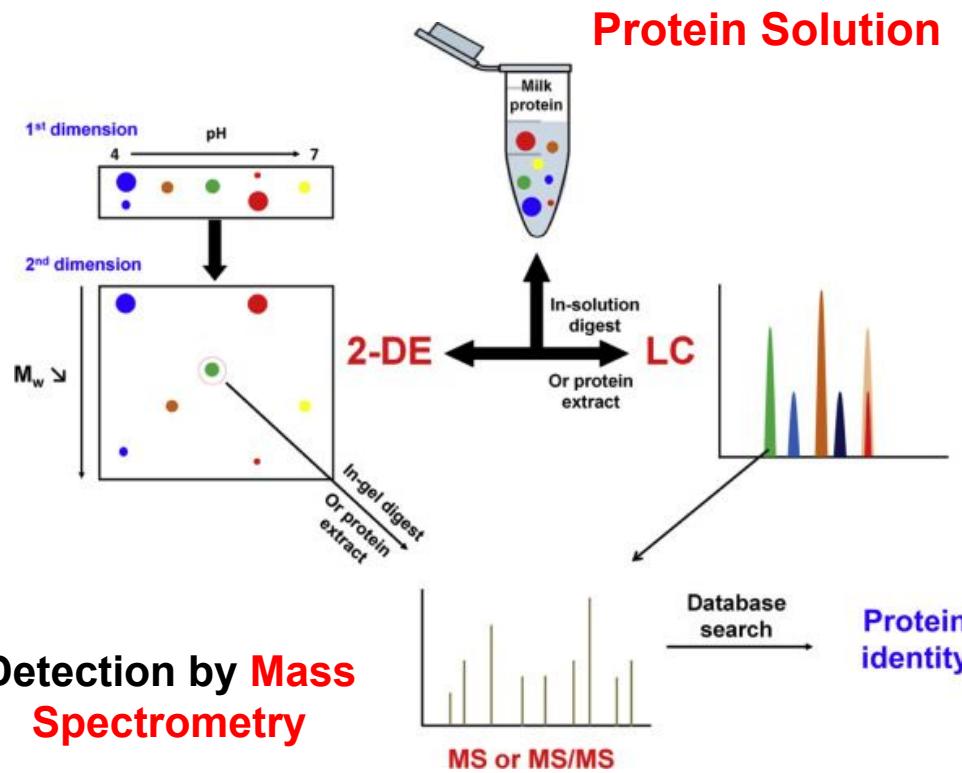
Molecular Weight  
of Protein

# Measurements of Gene Expression

- **mRNA-based Measurements**
  - **Northern Blot:** Based on Hybridization technique and fluorescent probe
  - **RT-PCR:** Based on PCR and Reverse Transcriptase (RT)
  - **RT-qPCR:** Based on PCR with fluorescent dye
  - **Microarray:** Based on Hybridization technique and fluorescent probe
  - **RNA-seq:** Based on NGS
- **Protein-based Measurements**
  - **Western Blot:** Based on antigen-antibody binding and PAGE
  - **Mass Spectrometry**
    - **Two-Dimensional Polyacrylamide Gel Electrophoresis (2-DE) -> MS**
    - **LC and HPLC -> MS**

# Proteomic Analysis based on Mass Spectrometry

Separate Protein by 2-DE (Charge and MW)

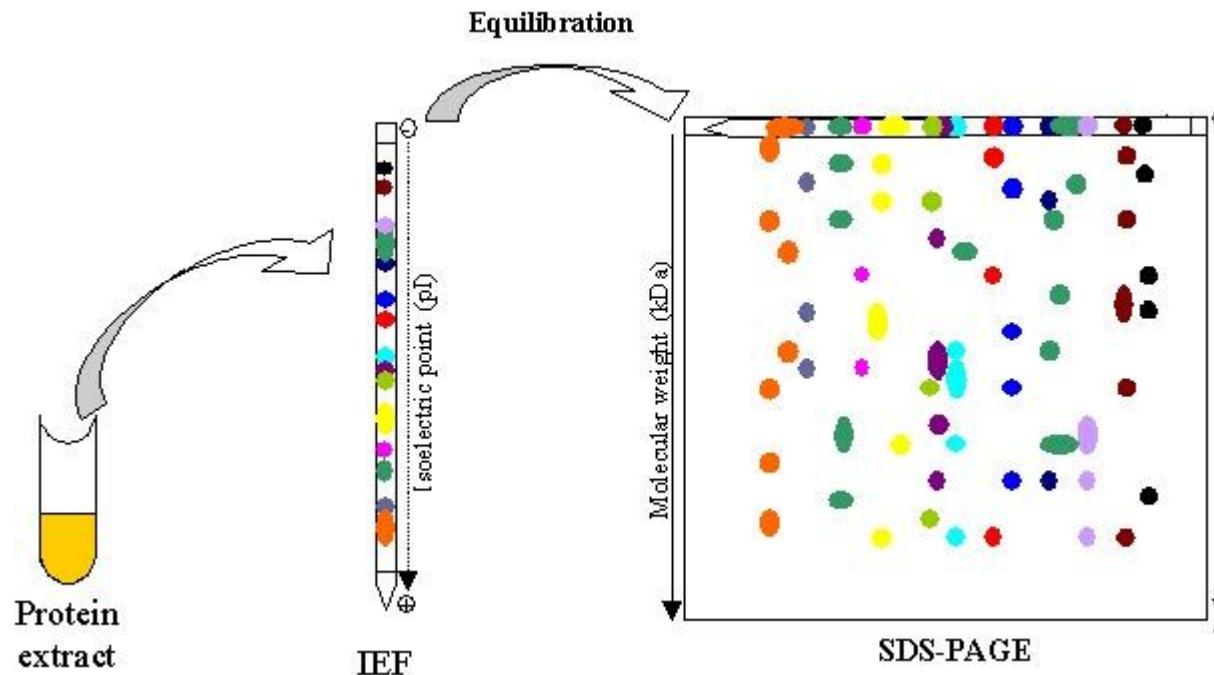


Separate Protein by Liquid Chromatography (LC)

Compare MS results with Protein Database

# 2-DE -> MS

- **2-Dimensional Gel Electrophoresis (2-DE)**



# First dimension: Isoelectric Focusing (IEF)

- IsoElectric Focusing (IEF)
- Use pH gradient and electric field to separate Protein by isoelectric point

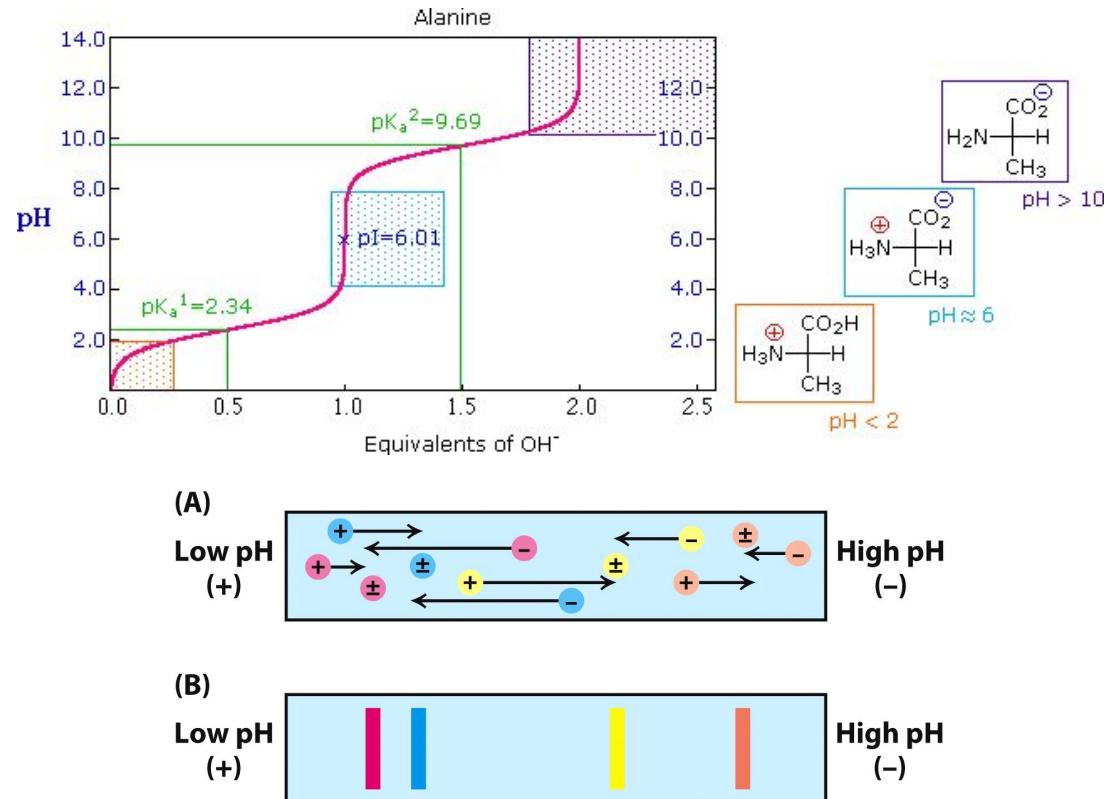
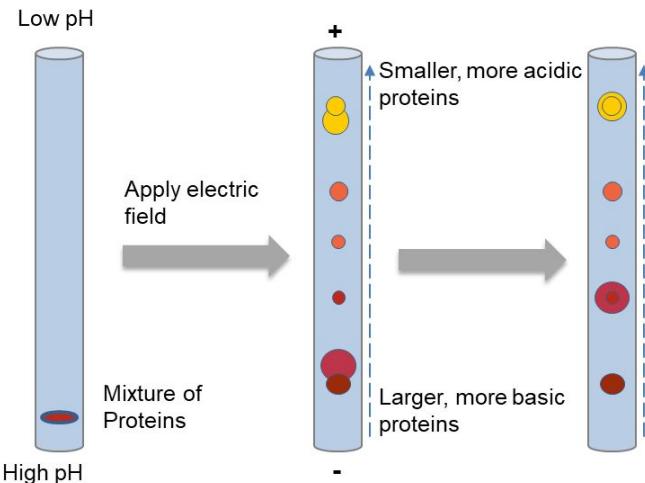
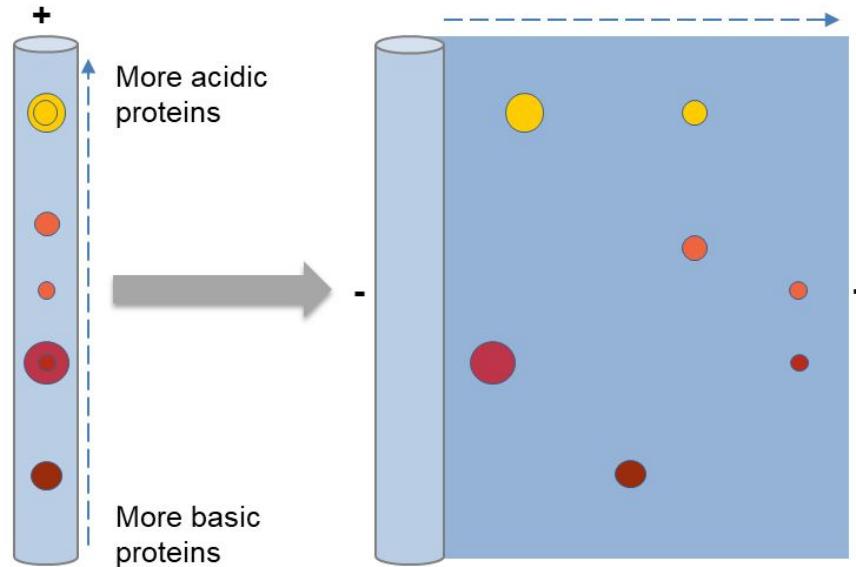


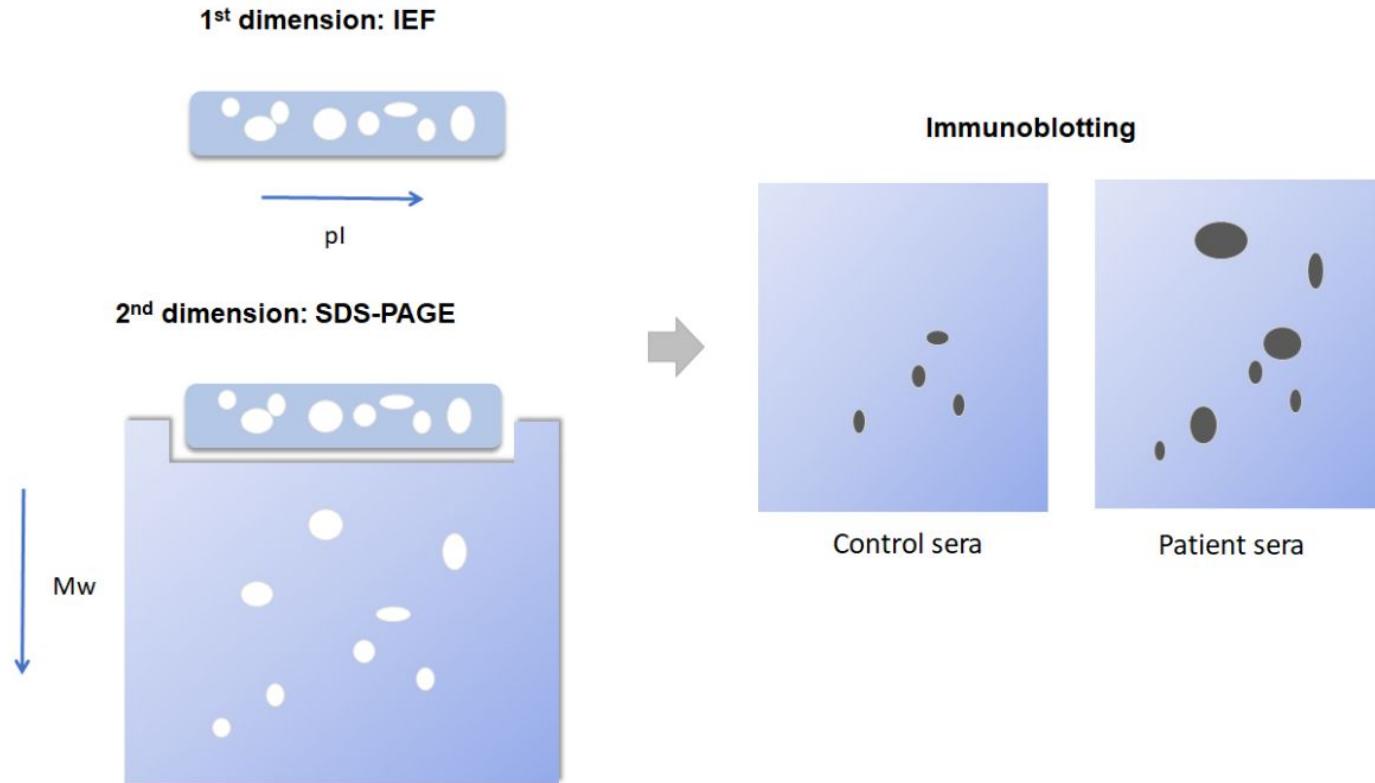
Figure 3.11  
Biochemistry, Seventh Edition  
© 2012 W. H. Freeman and Company

# Second dimension: Size Focusing

- SDS-PAGE
  - Separate Protein by **Size**

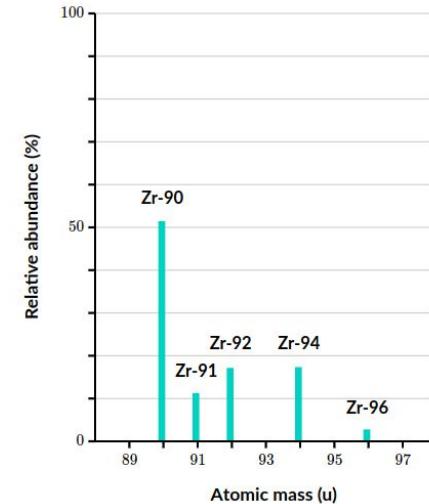
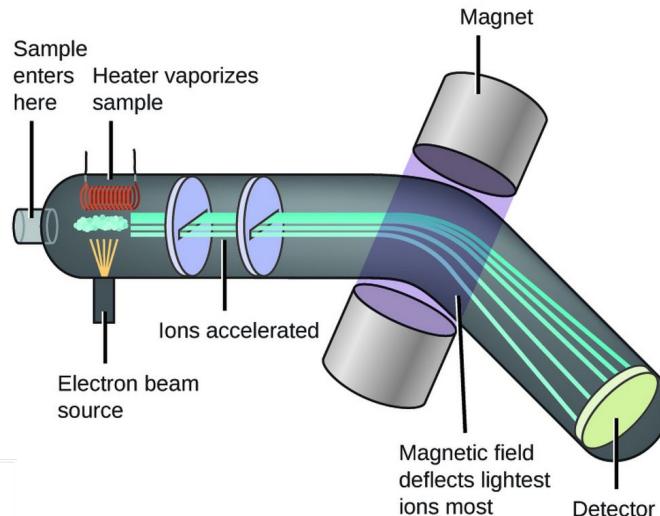
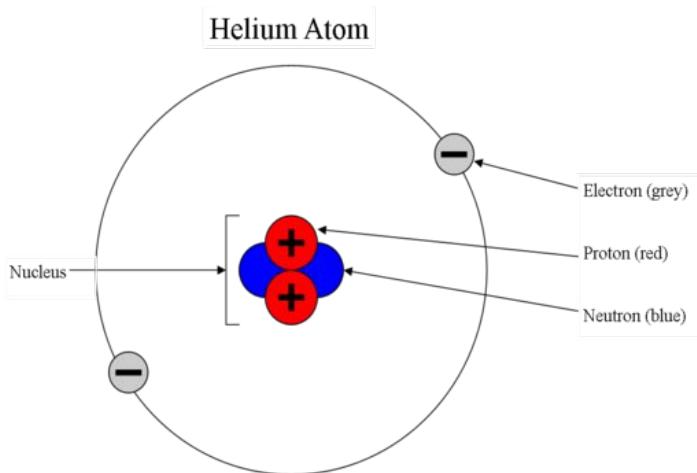


# 2-DE



# Mass Spectrometry (MS)

- Vaporize
- Ionize
- Accelerate
- Blend
- Detect



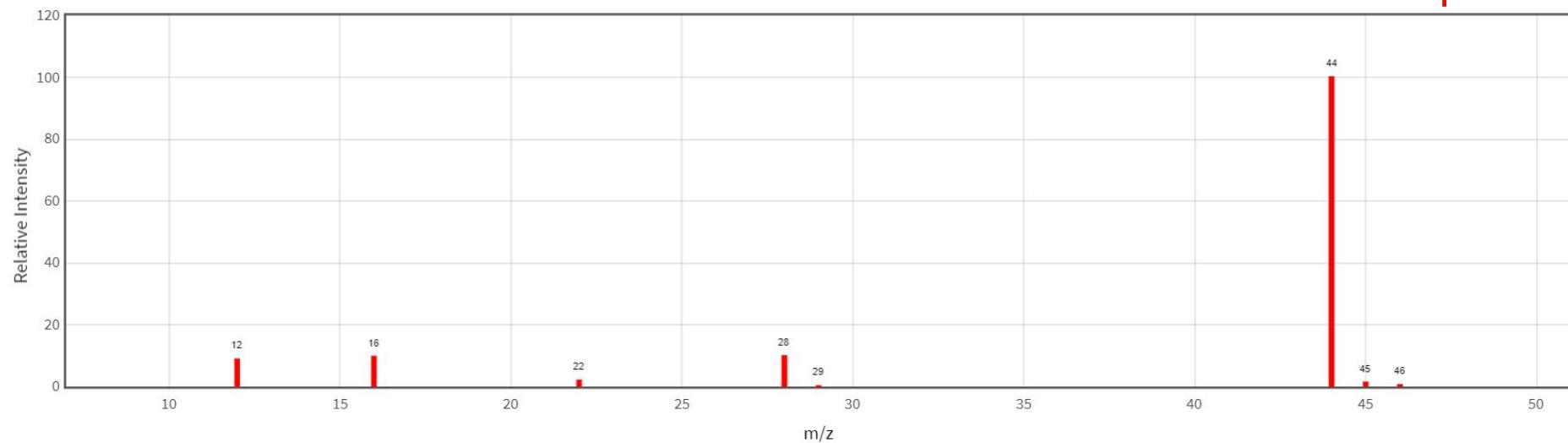
# MS output for Carbon Dioxide



116.3 pm

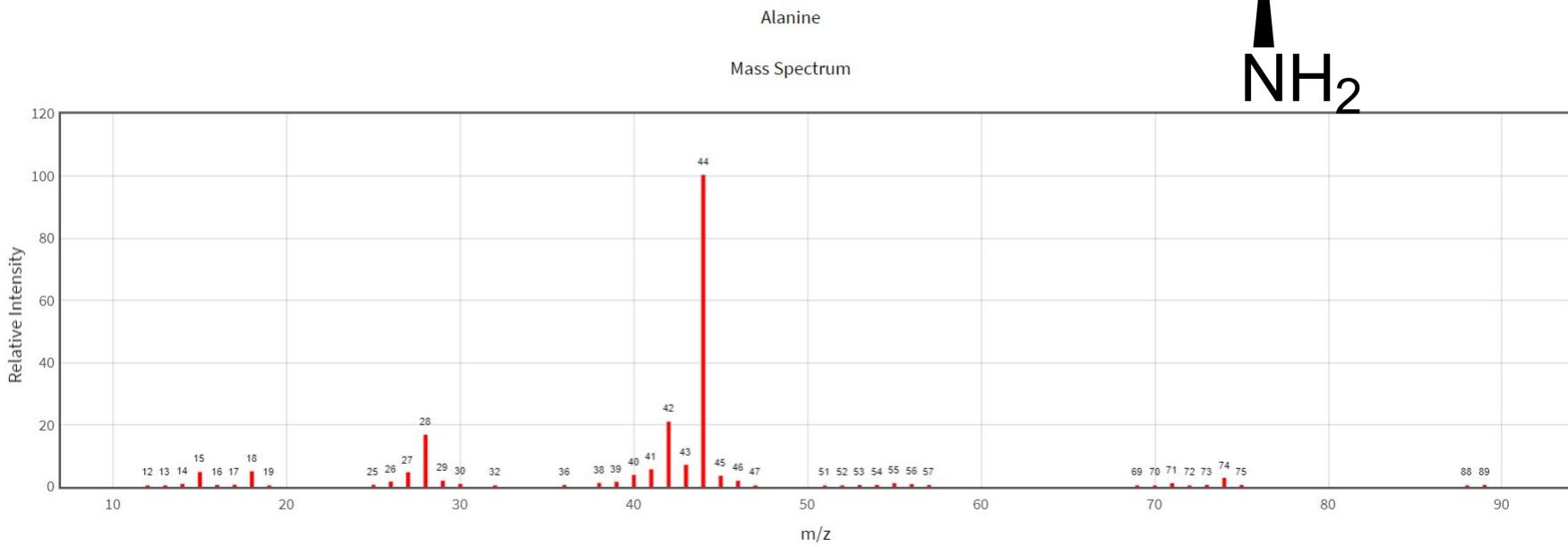
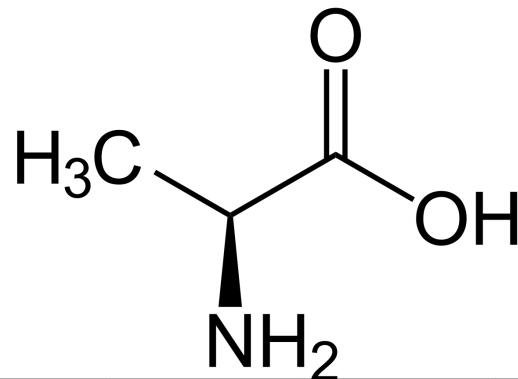
Carbon dioxide

Mass Spectrum

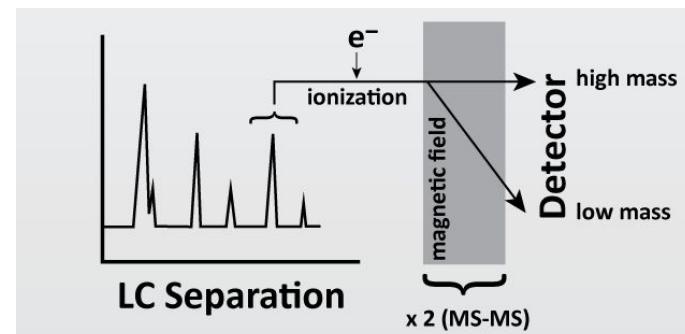
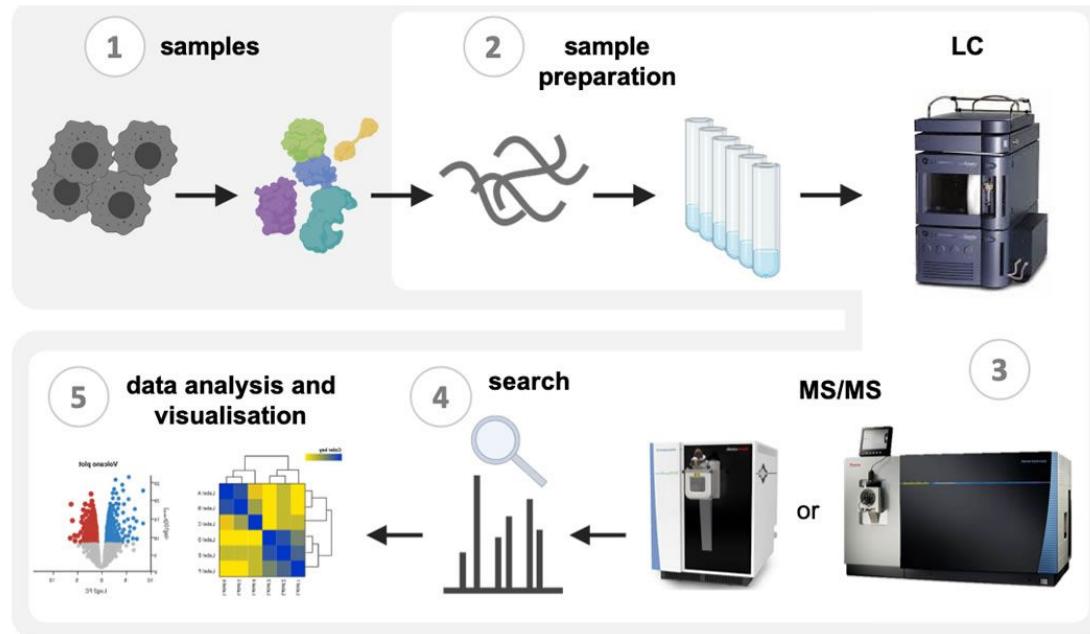


$\text{C} = 12, \text{O} = 16$

# MS output for Alanine



# Proteomic Analysis: LC → MS



Created with BioRender.com