Enzyme-based molecular techniques

Prof. Mamoun Ahram
Dr. Diala Abu Hassan
Summer 2019

Resources

- This lecture
- Cooper, pp. 127-129, 124-125, Ch. 5, pp.159-162, 166-171

What is DNA sequencing?

- DNA sequencing is the process of determining the exact order of nucleotides in a genome.
- Importance:
  - Identification of genes and their localization
  - Identification of protein structure and function
  - Identification of DNA mutations
  - Genetic variations among individuals in health and disease
  - Prediction of disease-susceptibility and treatment efficiency
  - Evolutionary conservation among organisms

DNA sequencing of organism genome

- Viruses and prokaryotes first
- Human mitochondrial DNA
- The first eukaryotic genome sequenced was that of yeast, Saccharomyces cerevisiae.
- The genome of a multicellular organism, the nematode Caenorhabditis elegans.
- Determination of the base sequence in the human genome was initiated in 1990 and completed in May 2006 via the Human Genome Project

SPECIES	BASE PAIRS (estimated)	GENES (estimated)	CHROMOSOMES
Human *(Homo sapiens)*	3.2 billion	~ 25,000	46
Mouse *(Mus musculus)*	2.6 billion	~ 25,000	40
Fruit Fly *(Drosophilia melanogaster)*	137 million	13,000	8
Roundworm *(Caenorhabditis elegans)*	97 million	19,000	12
Yeast *(Saccharomyces cerevisia)*	12.1 million	6,000	32
Bacteria (Escherichia coli)	4.6 million	3,200	1
Bacteria (H. influenzae)	1.8 million	1,700	1

Nucleotides per genomes

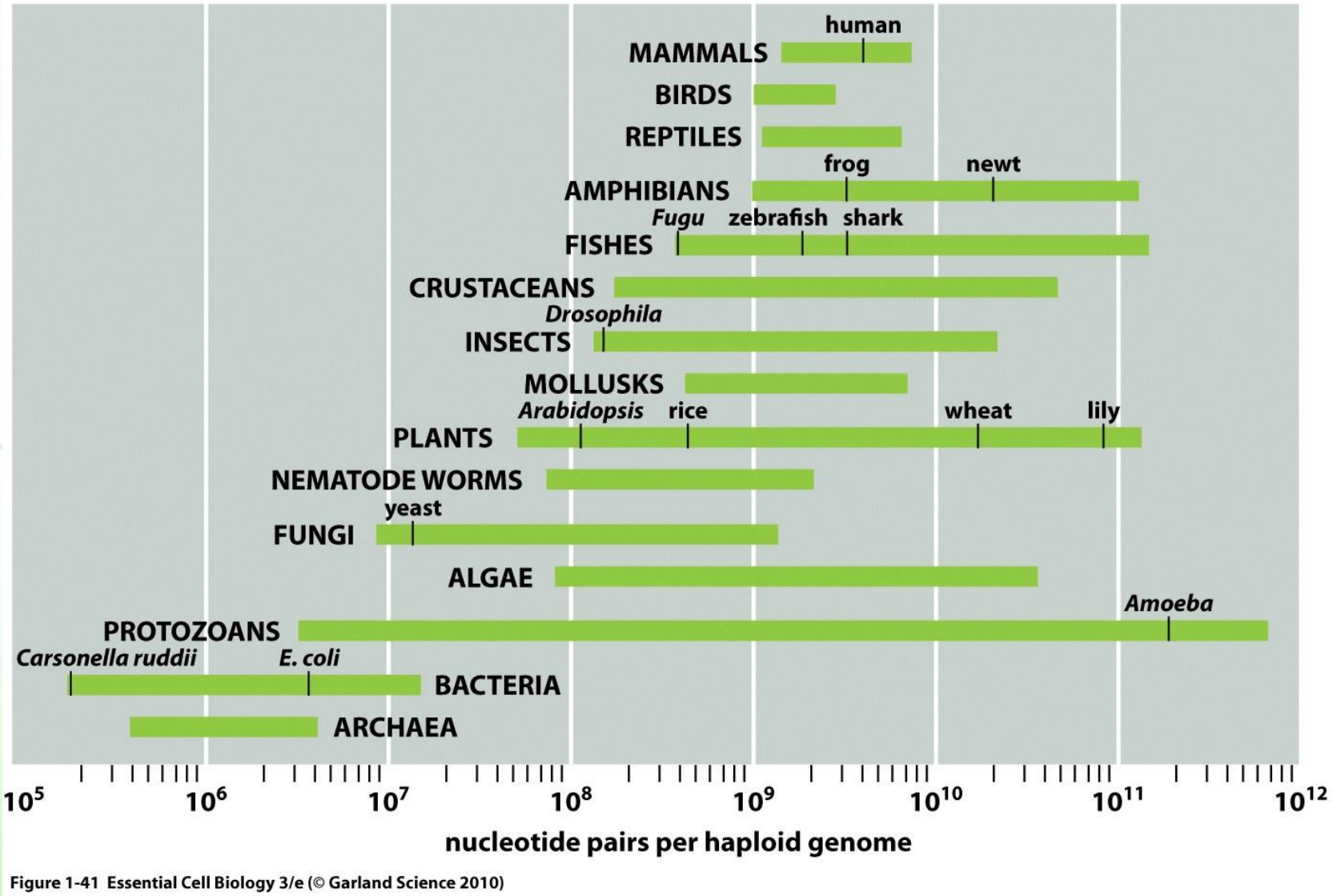
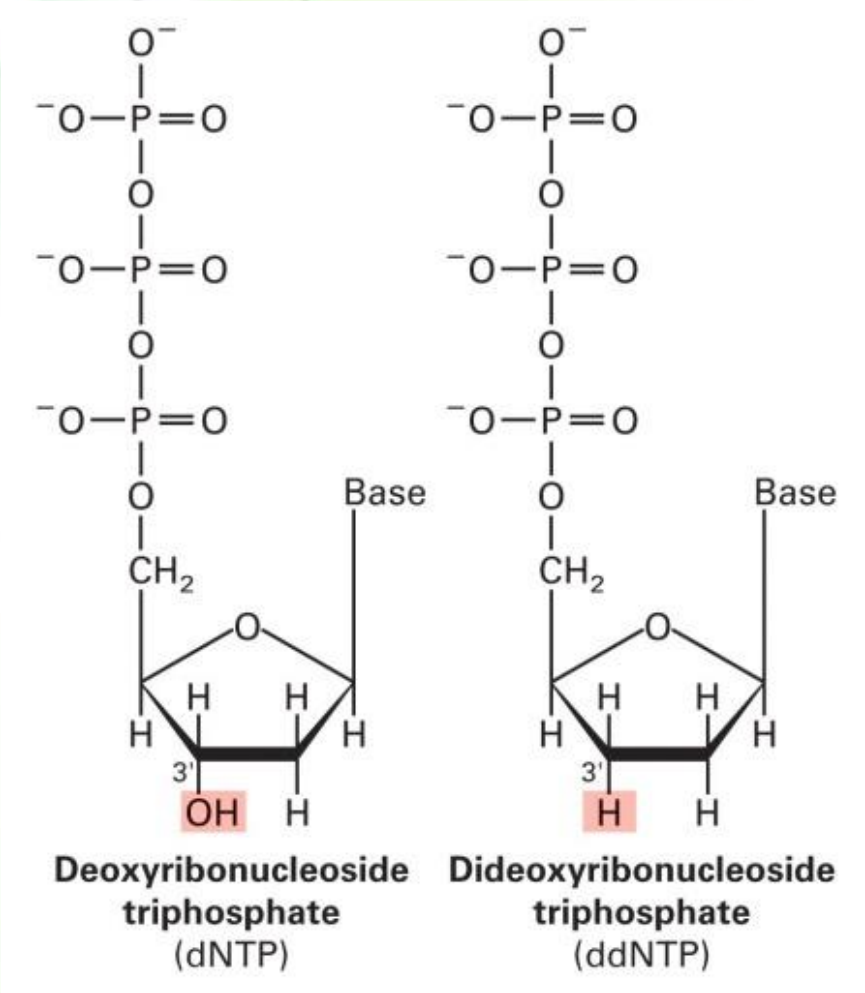


Figure 1-41 Essential Cell Biology 3/e (© Garland Science 2010)

Method of DNA sequencing

- The most popular method is based on premature termination of DNA synthesis by dideoxynucleotides.

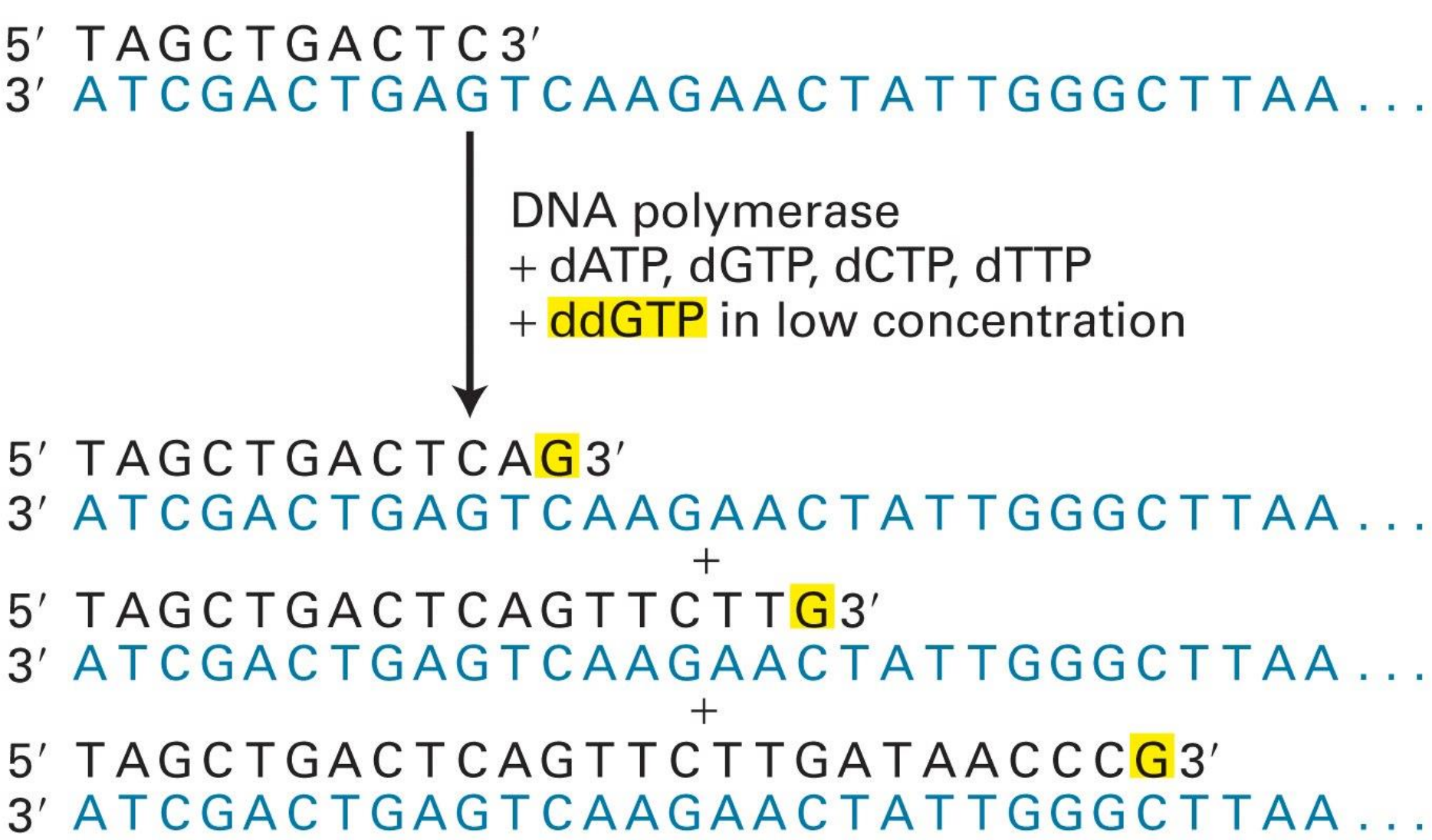


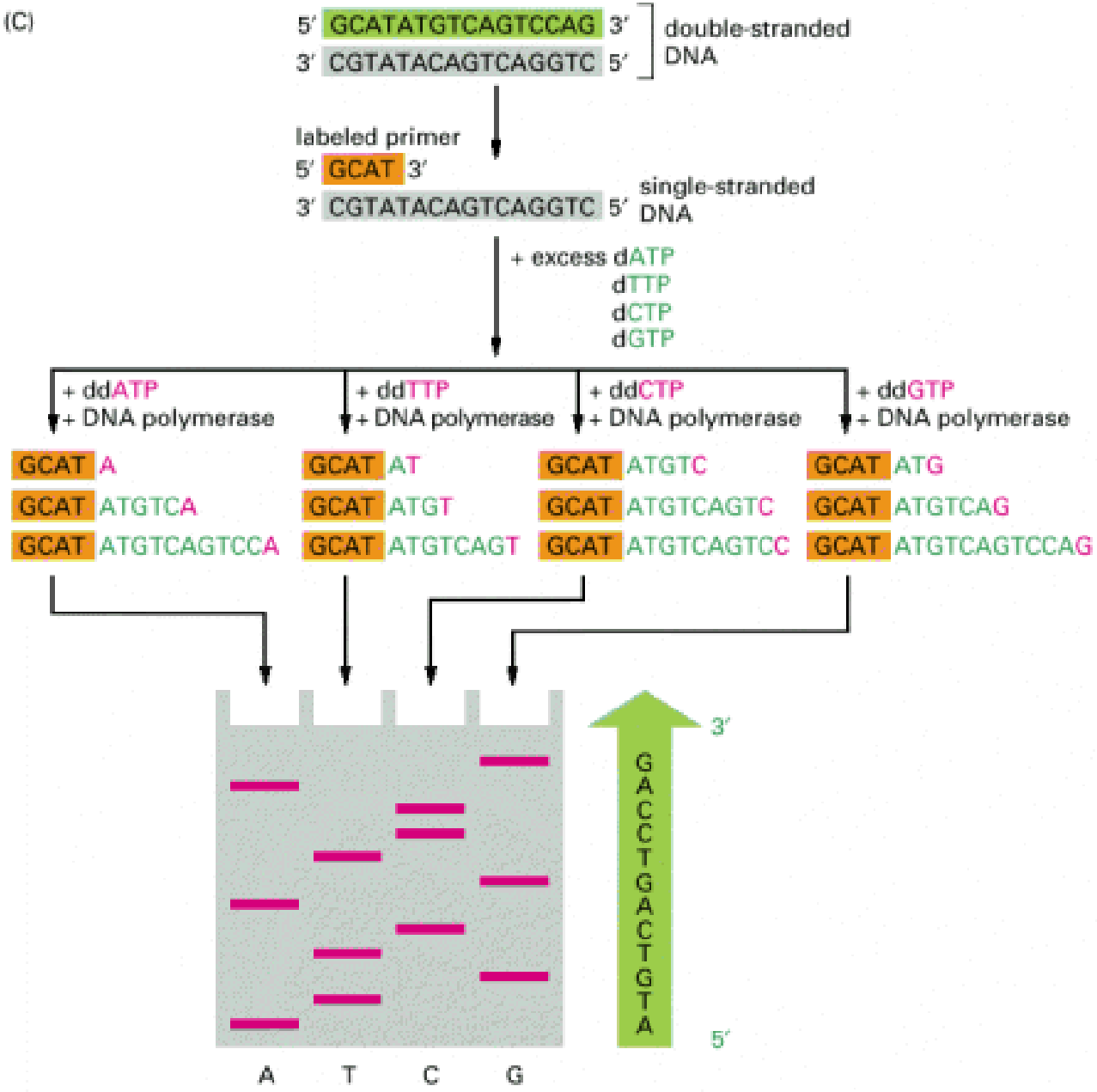
The process...

- DNA synthesis is initiated from a primer that has been labeled with a radioisotope
- Four separate reactions are run, each including deoxynucleotides plus one dideoxynucleotide (either A, C, G, or T)
- Incorporation of a dideoxynucleotide stops further DNA synthesis because no 3 hydroxyl group is available for addition of the next nucleotide

Generation of fragments

- A series of labeled DNA molecules are generated, each terminated by the dideoxynucleotide in each reaction
- These fragments of DNA are then separated according to size by gel electrophoresis and detected by exposure of the gel to X-ray film
- The size of each fragment is determined by its terminal dideoxynucleotide, so the DNA sequence corresponds to the order of fragments read from the gel

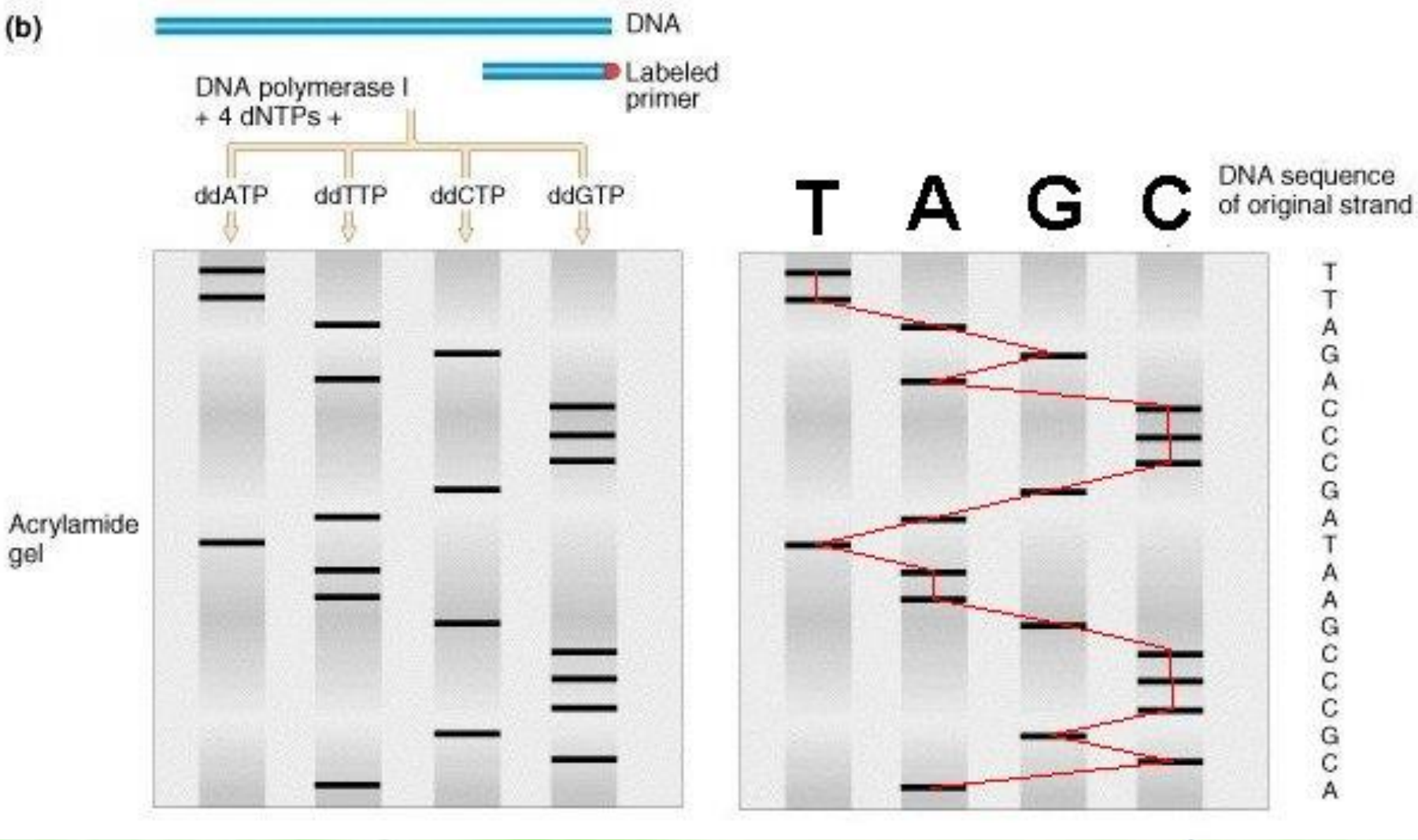

5′ TAGCTGACTC 3′
3′ ATCGACTGAGTCAAGAACTATTGGGCTTAA . . .
DNA polymerase
+ dATP, dGTP, dCTP, dTTP
+ ddGTP in low concentration
5′ TAGCTGACTCAG 3′
3′ ATCGACTGAGTCAAGAACTATTGGGCTTAA . . .
+
5′ TAGCTGACTCAGTTCTTG 3′
3′ ATCGACTGAGTCAAGAACTATTGGGCTTAA . . .
+
5′ TAGCTGACTCAGTTCTTGATAACCCG 3′
3′ ATCGACTGAGTCAAGAACTATTGGGCTTAA . . .

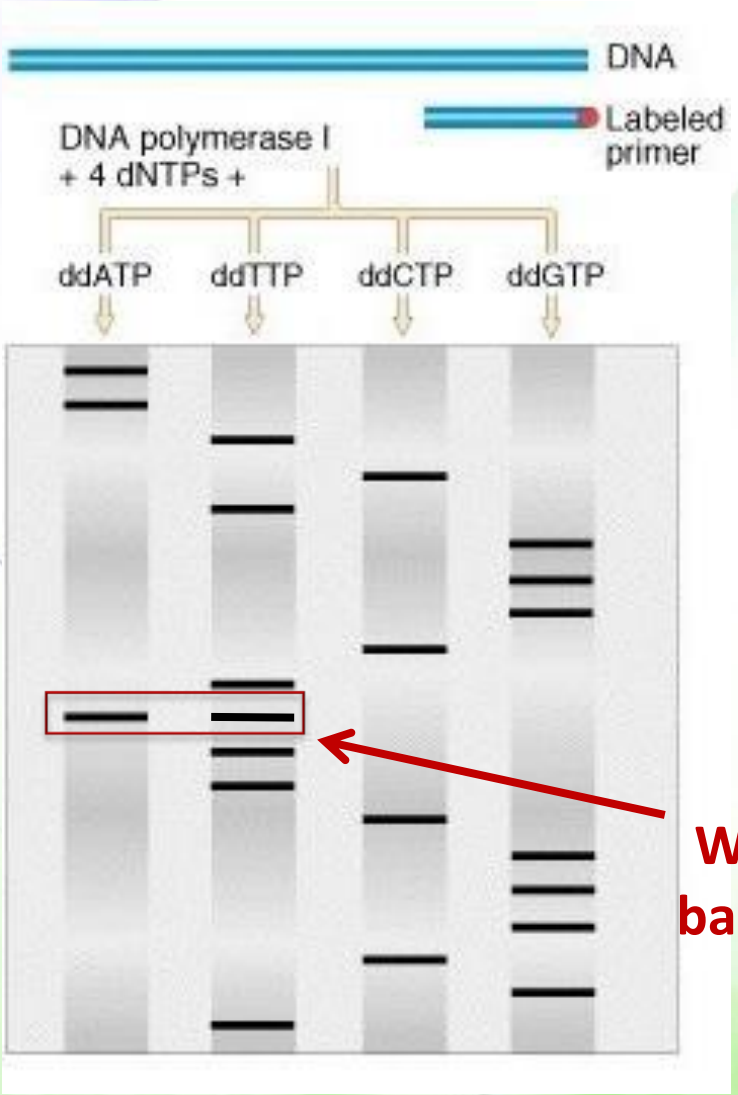


DNA sequence reading directly from the bottom of the gel upward, is

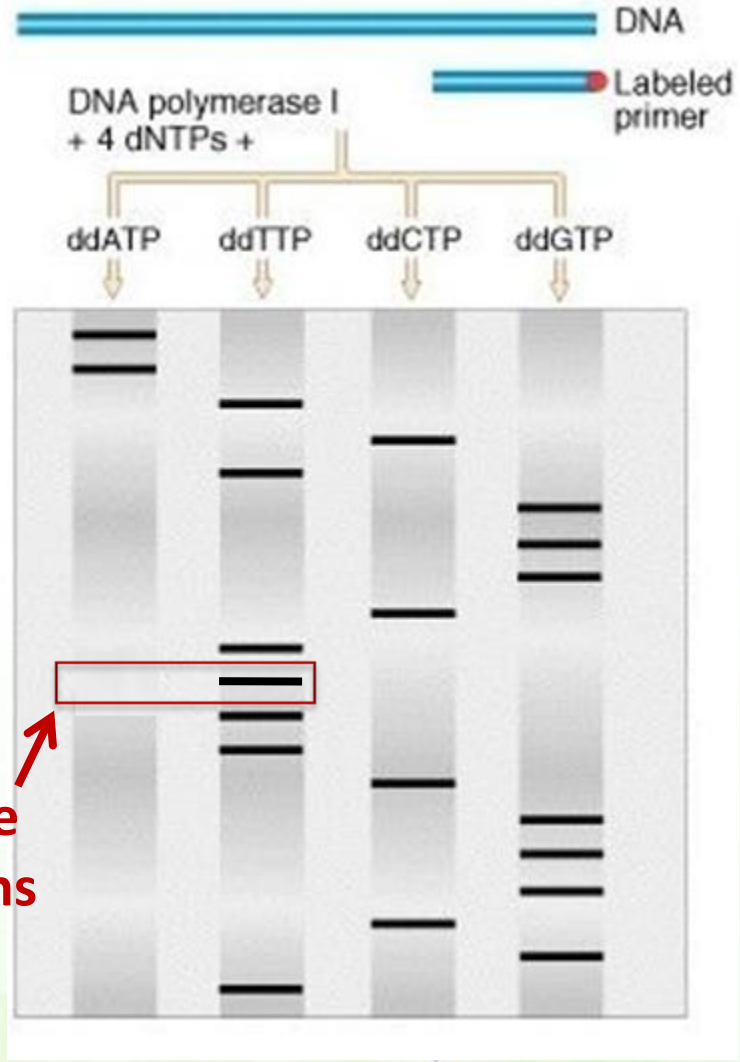
ATGTCAGTCCAG

1 12


(b)
DNA
Labeled primer
DNA polymerase I
+ 4 dNTPs +
ddATP
ddTTP
ddCTP
ddGTP
Acrylamide gel
T A G C
DNA sequence of original strand
T T A G A C C C G A T A A G C C C G C A

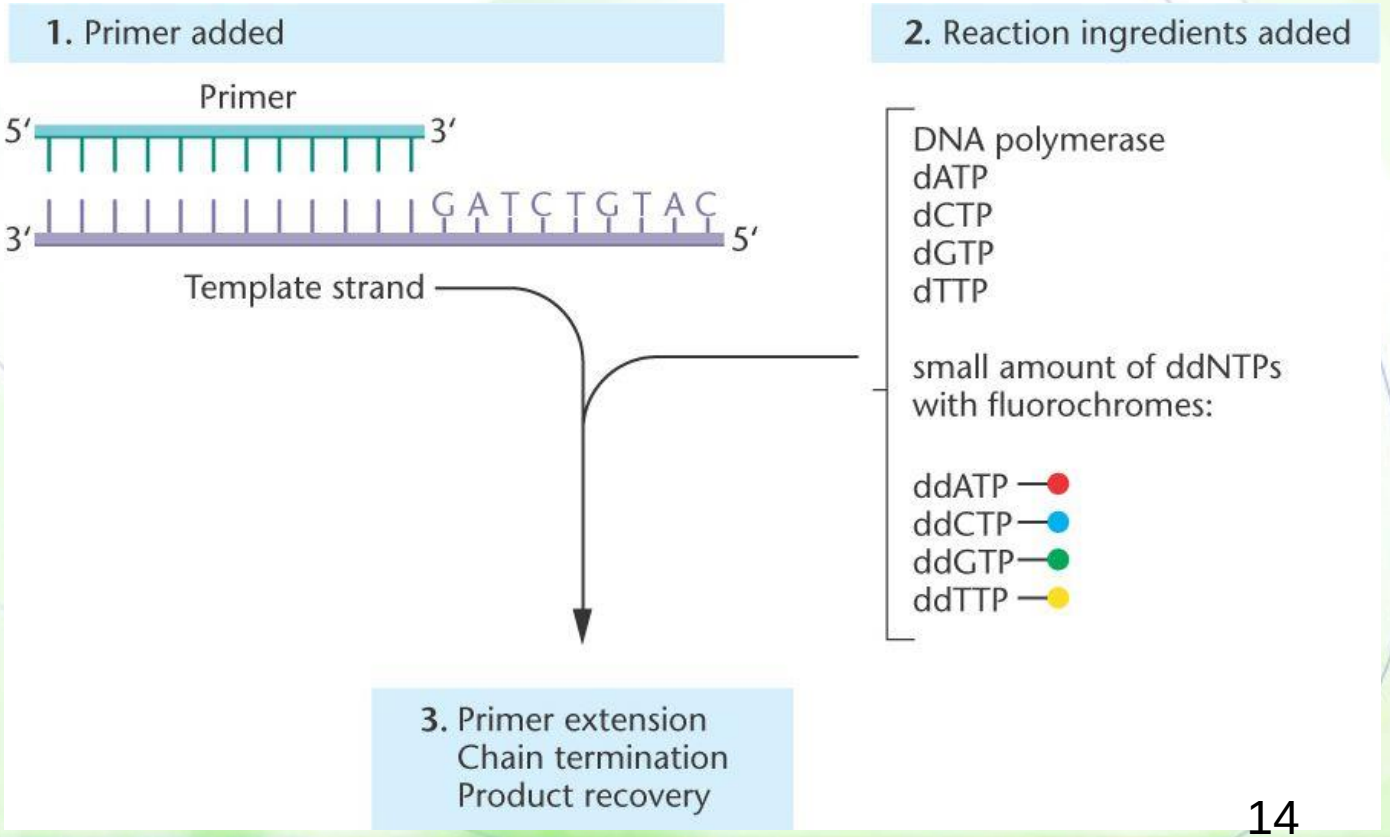


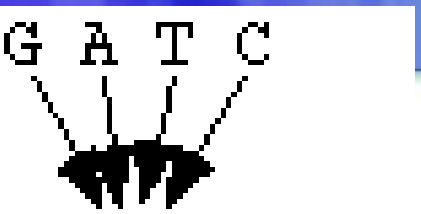
What do the band patterns mean?

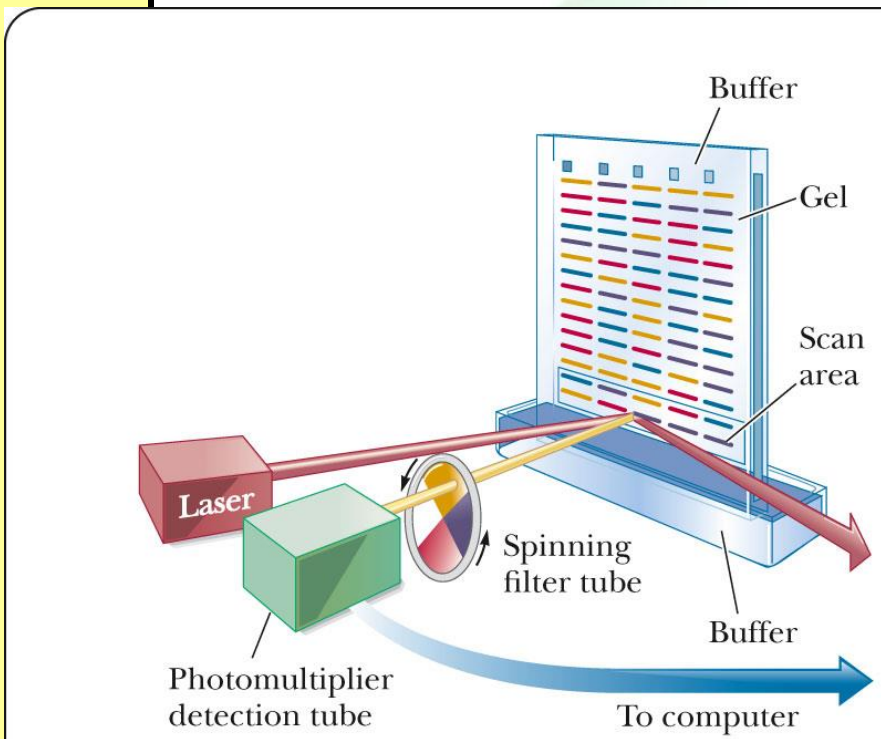


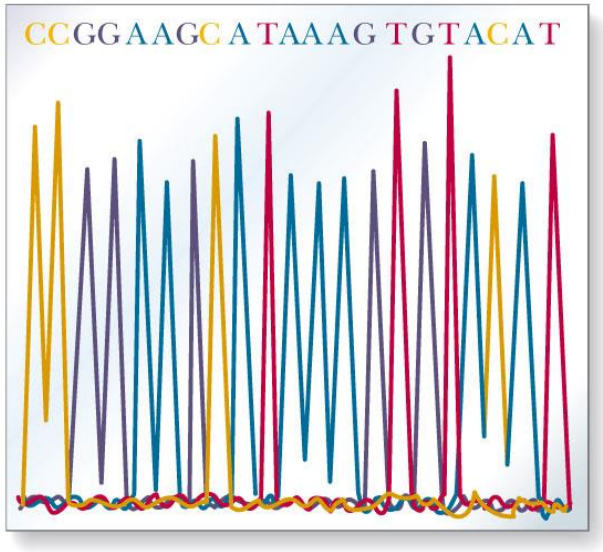
Fluorescence-based DNA sequencing

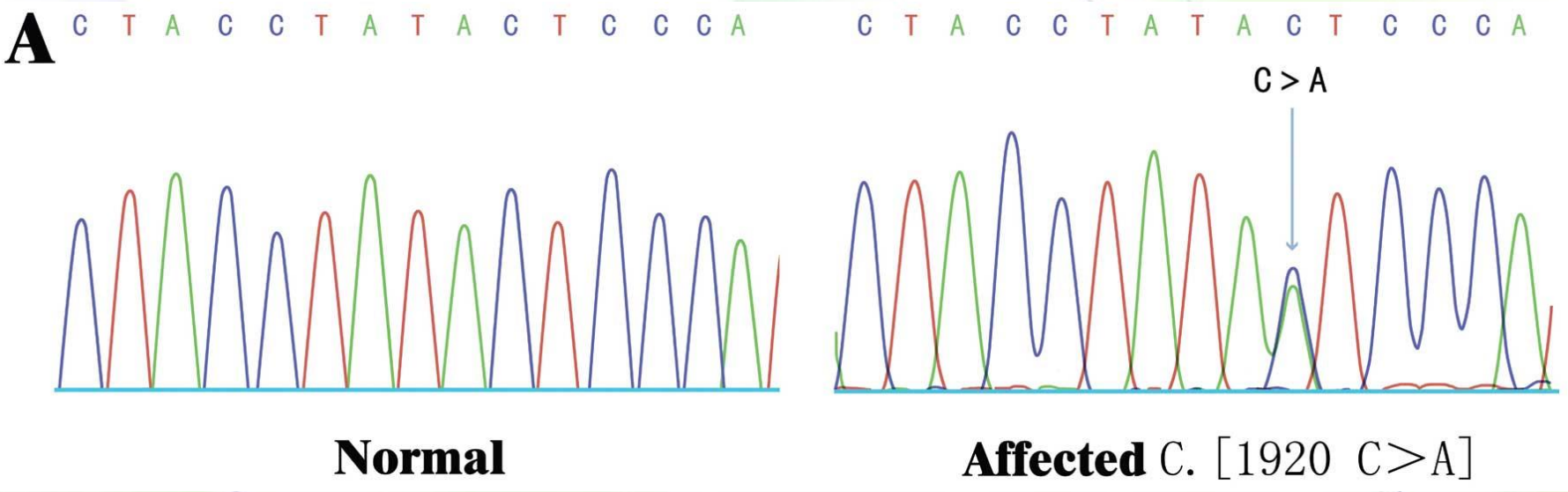
- Reactions include the four deoxynucleotides plus the four dideoxynucleotides in the same reaction with ***each ddNTP labeled with a unique fluorescent tag***.




G A T C


Buffer
Gel
Scan area
Laser
Spinning filter tube
Buffer
Photomultiplier detection tube
To computer

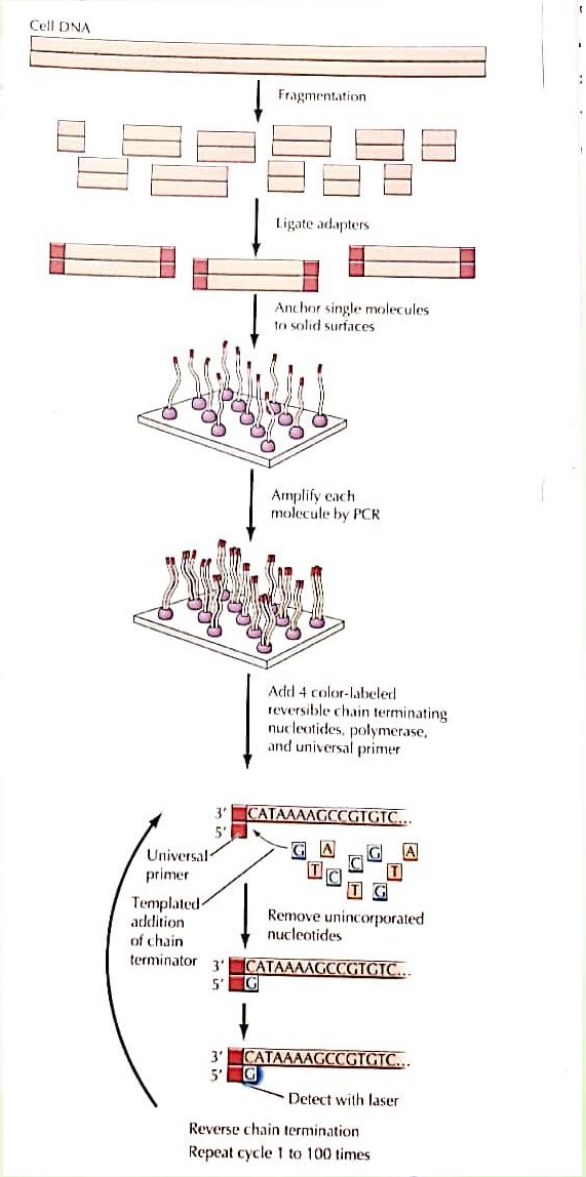

CCGGAAGCATAAAGTGTACAT



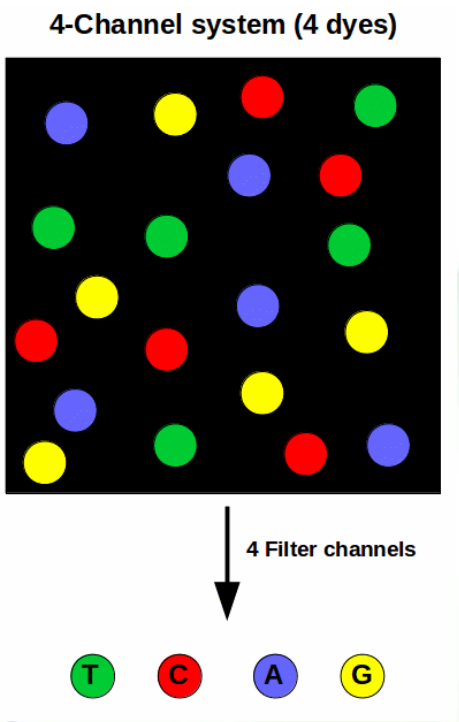
What does it mean?

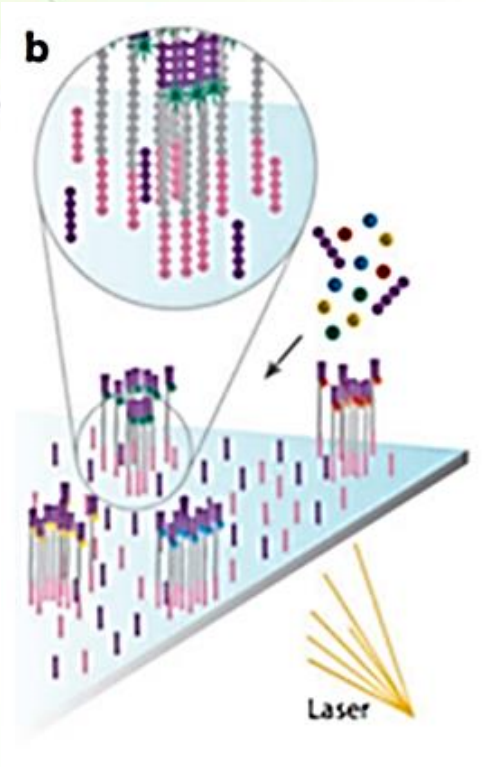
Next-generation sequencing

- Cellular DNA is fragmented.
- DNA adapters are added to ends of each DNA fragment.
- Each DNA fragment is attached to a solid surface and amplified like PCR using primers that anneal to the adapter sequences.
- Four-color nucleotides with terminating ends are added.
- Incorporation of each single nucleotide is detected.
- Unincorporated nucleotides are removed.
- The cycle is repeated.

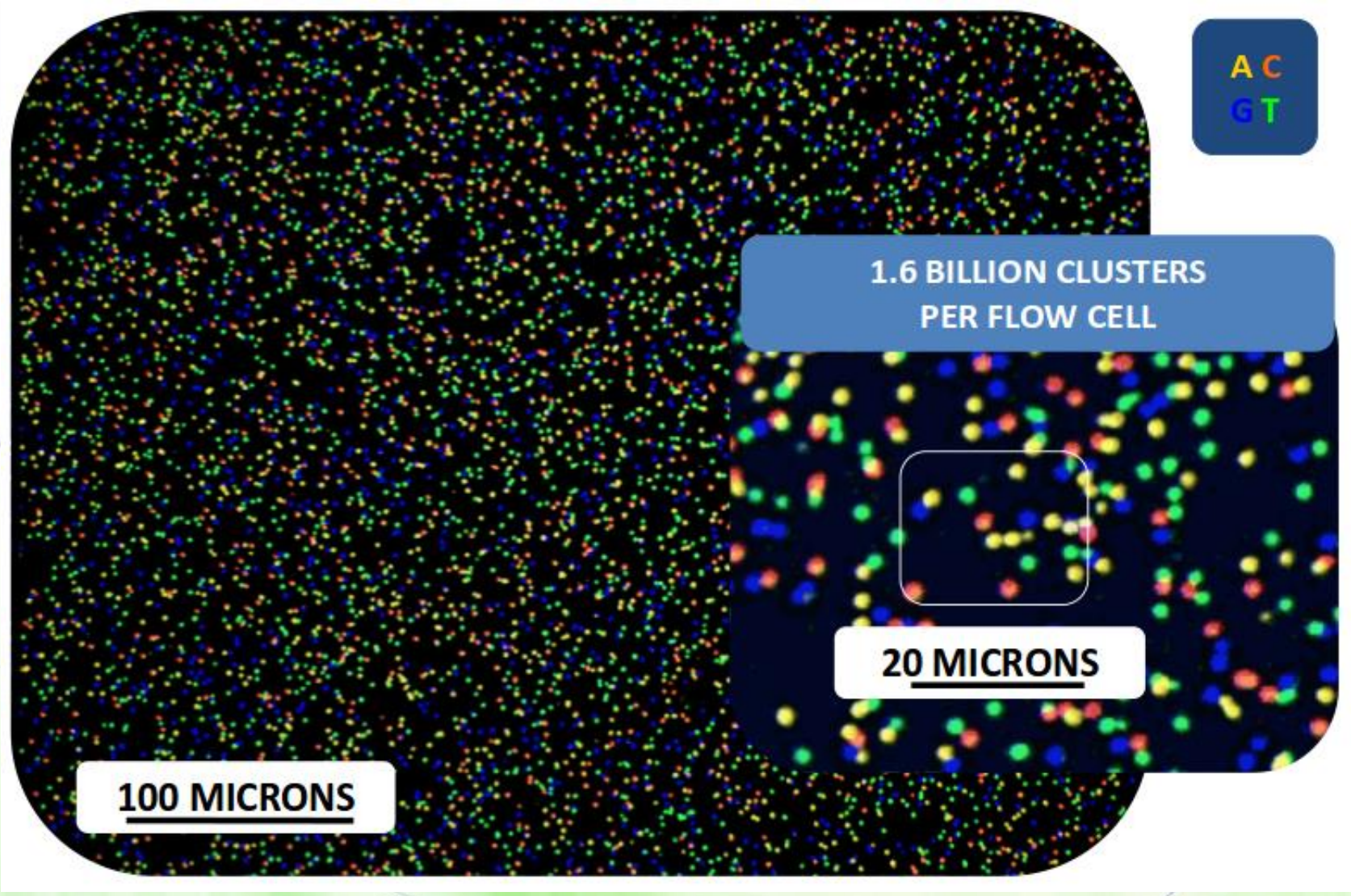


The detection





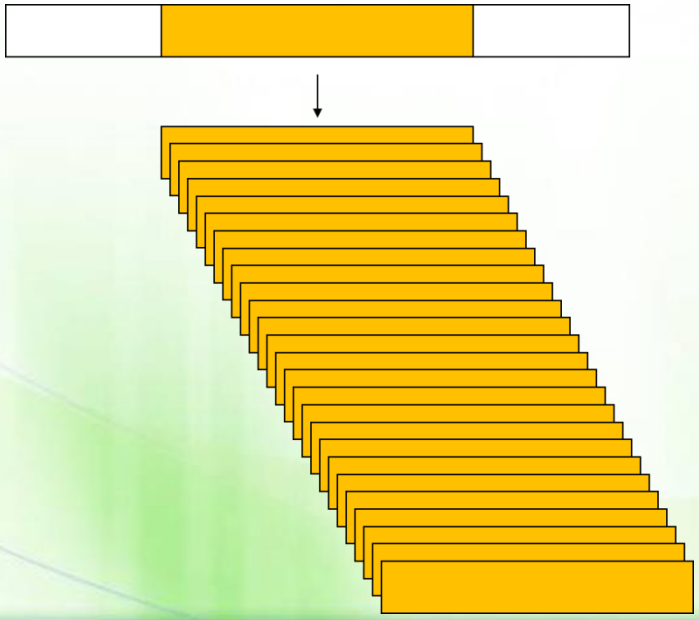
A real look


A C
G T
1.6 BILLION CLUSTERS
PER FLOW CELL
20 MICRONS
100 MICRONS

Polymerase chain reaction (PCR)

Polymerase Chain Reaction

- Polymerase chain reaction (PCR) allows the DNA from a selected region of a genome to be amplified a billionfold, effectively "purifying" this DNA away from the remainder of the genome.
- It s extremely sensitive; it can detect a single DNA molecule in a sample.

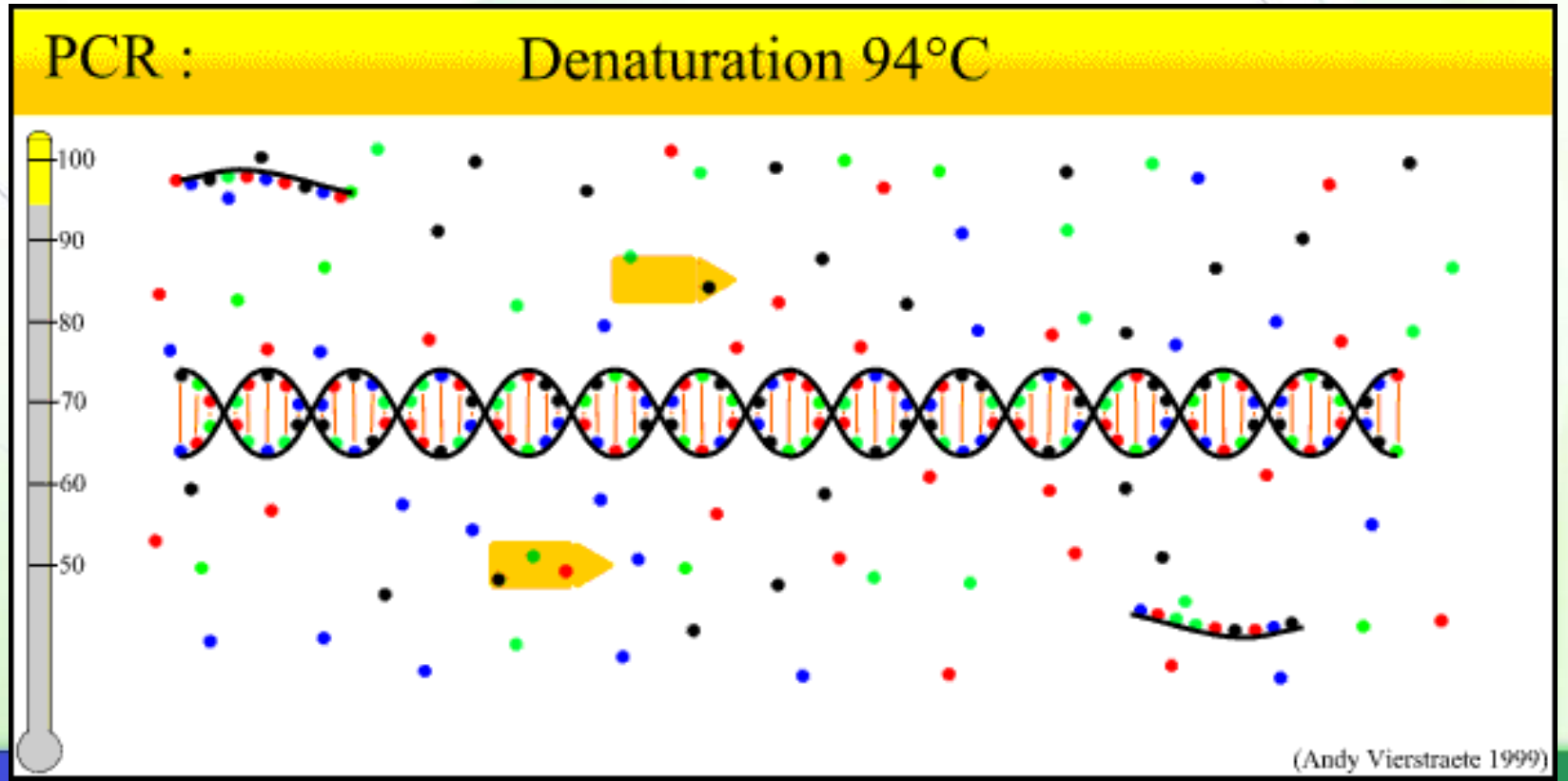


Components of PCR reaction

- The DNA template
- A pair of primers
  - The 15-25 nucleotides-long primers should surround the target sequence.
- All four deoxyribonucleoside triphosphates
- A heat-stable DNA polymerase

The PCR *cycles*

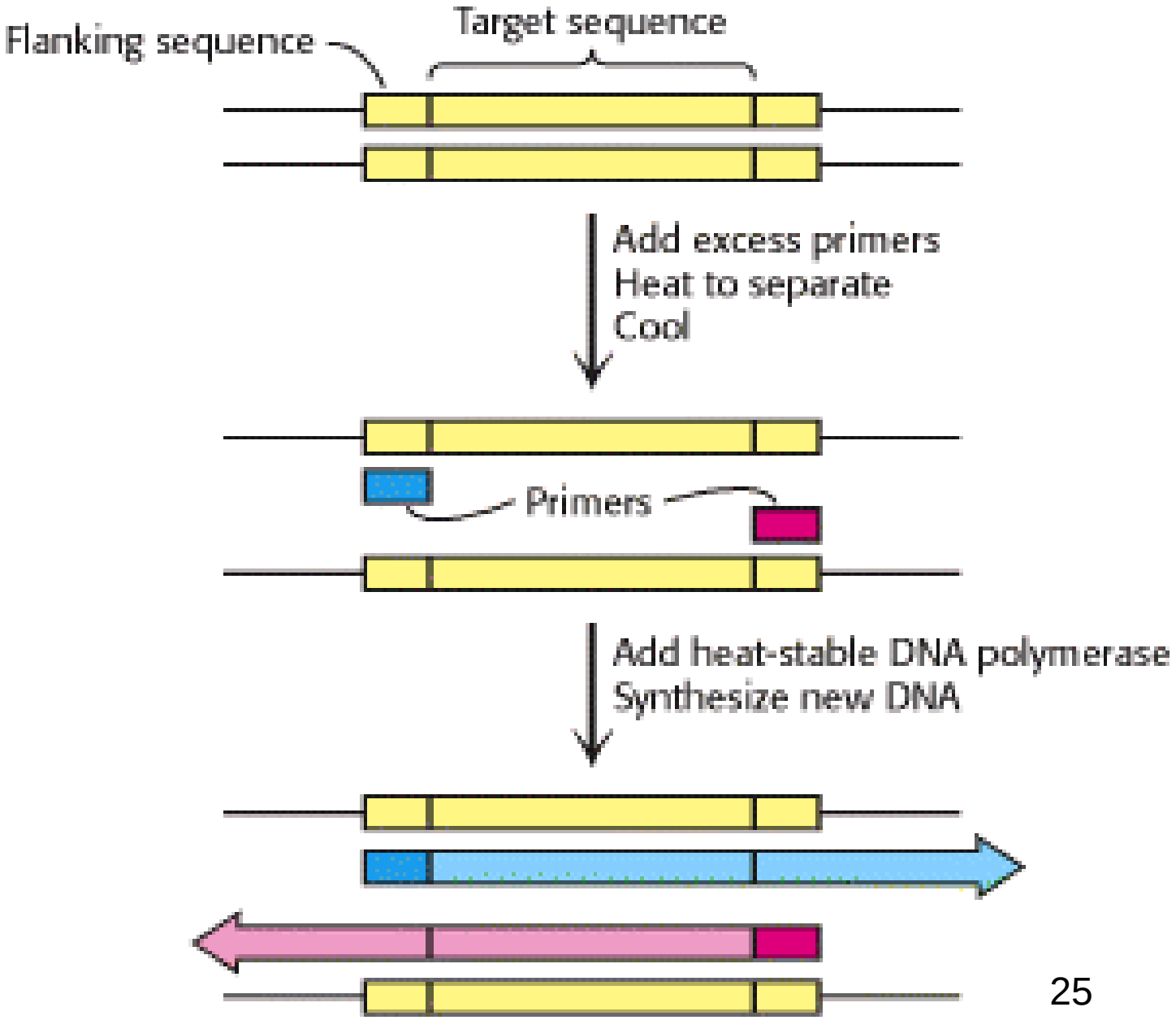
- Denaturation (at 95°C): DNA is denatured into single-stranded molecules.
- Reannealing (50°C to 70°C): the primers anneal to the DNA.
- DNA synthesis (at 72°C): optimal for the polymerase.

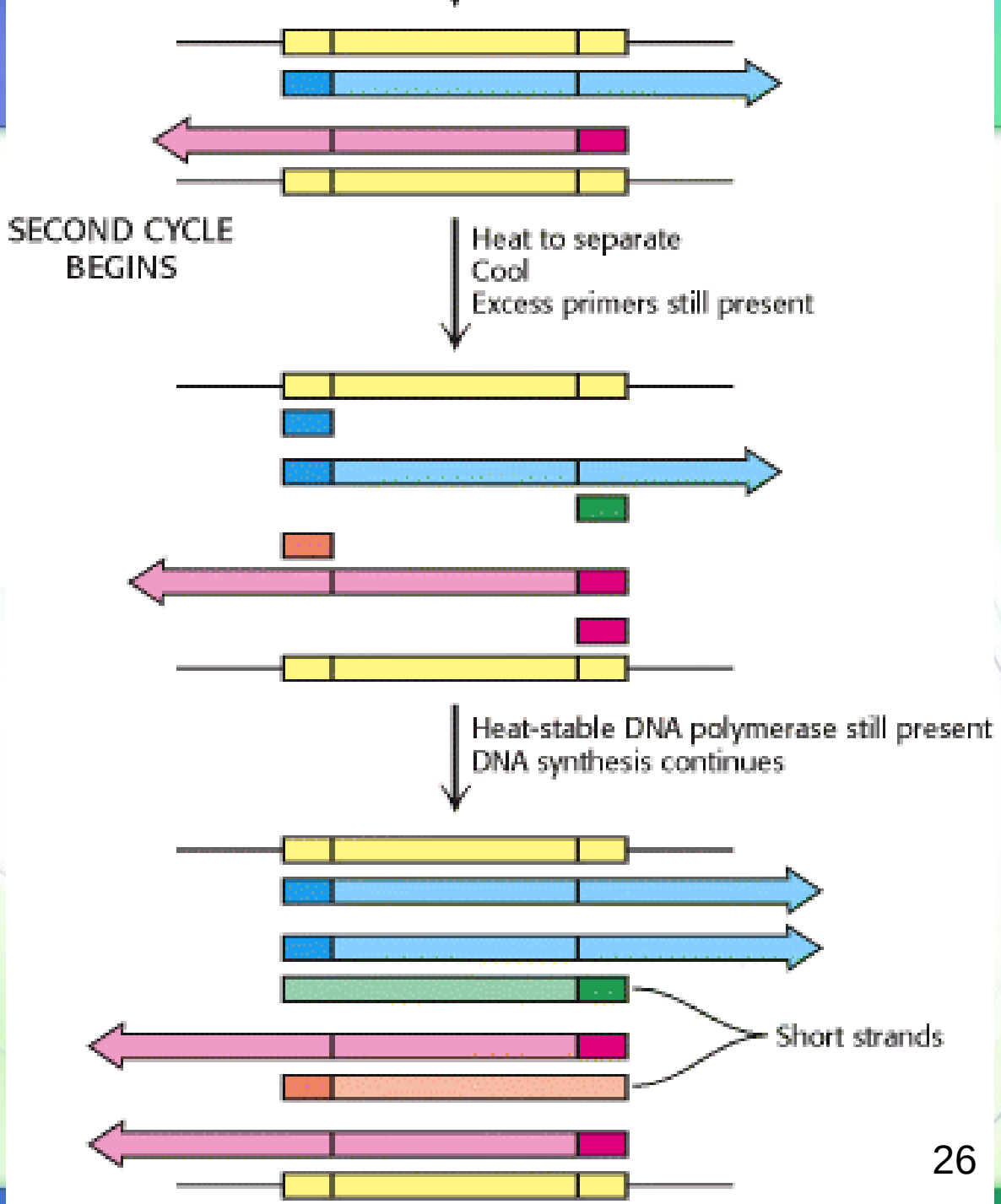


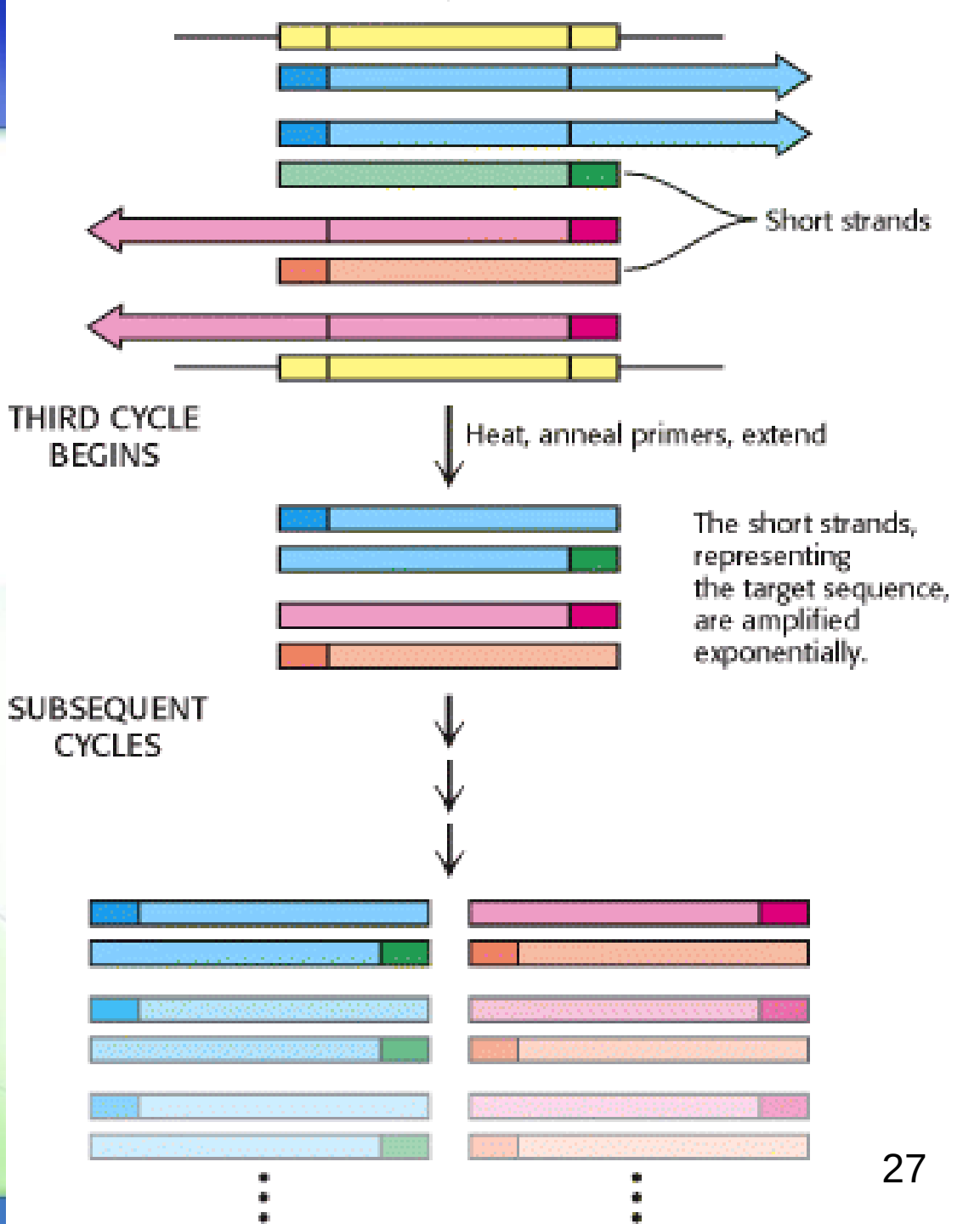
The DNA polymerase

- Suitably heat-stable DNA polymerases have been obtained from microorganisms whose natural habitat is hot springs.
- For example, the widely used Taq DNA polymerase is obtained from a thermophilic bacterium, Thermus aquaticus, and is thermostable up to 95°C.

FIRST CYCLE BEGINS




SECOND CYCLE
BEGINS
Heat to separate
Cool
Excess primers still present
Heat-stable DNA polymerase still present
DNA synthesis continues
Short strands

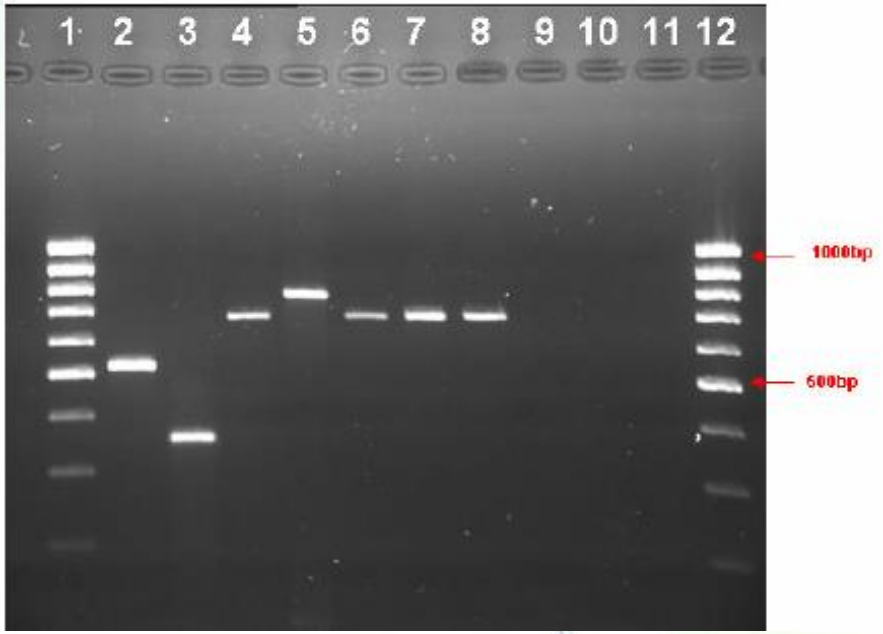

Short strands
THIRD CYCLE BEGINS
Heat, anneal primers, extend
The short strands, representing the target sequence, are amplified exponentially.
SUBSEQUENT CYCLES

PCR cycles

- 20-30 cycles of reaction are required for DNA amplification.
  - the products of each cycle serving as the DNA templates for the next-hence the term polymerase "chain reaction“.
- Every cycle doubles the amount of DNA.
- After 30 cycles, there will be over 250 million short products derived from each starting molecule.

Detection of DNA fragments

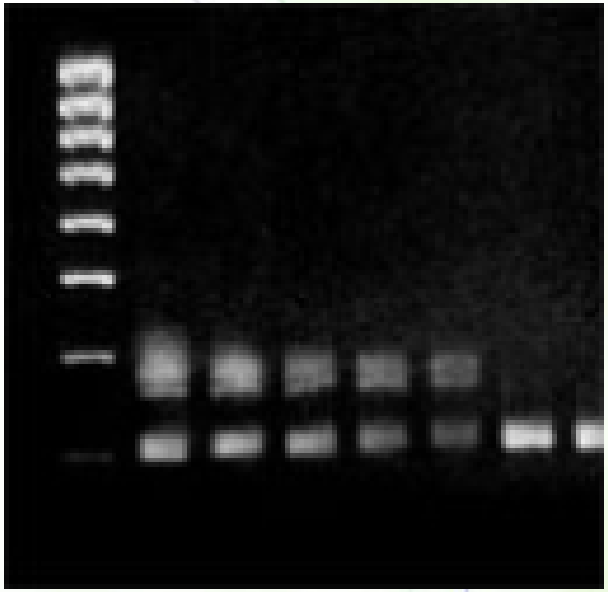
- This DNA fragment can be easily visualized as a discrete band of a specific size by agarose gel electrophoresis.



Importance of primers

- The specificity of amplification depends on the specificity of the primers to not recognize and bind to sequences other than the intended target DNA sequences
- How can you prevent it?
- How can you take advantage of it?

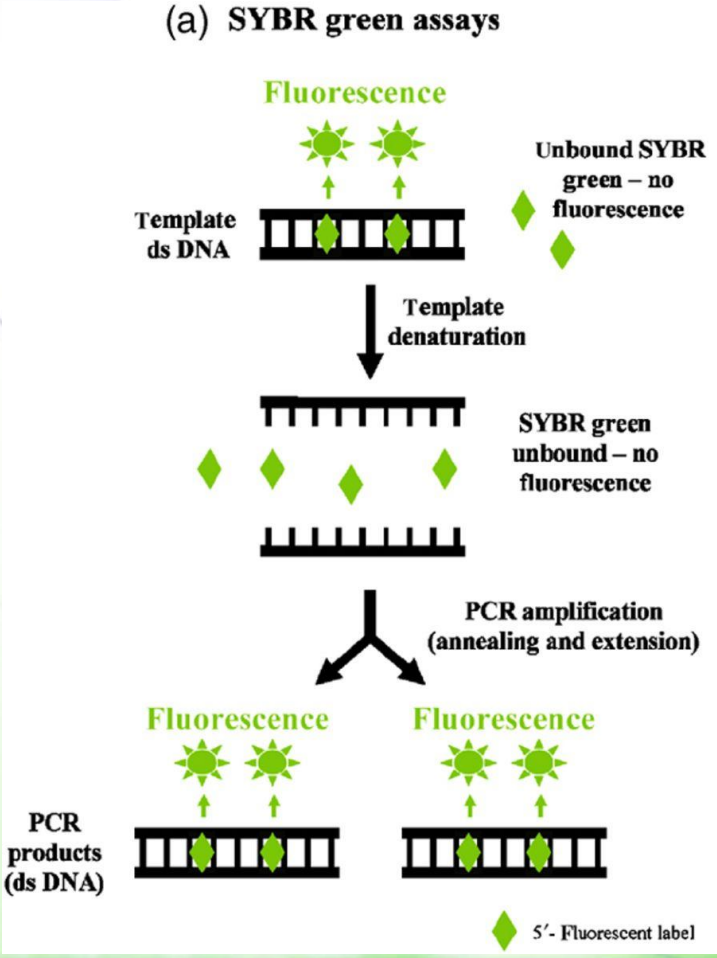
Annealing temperature



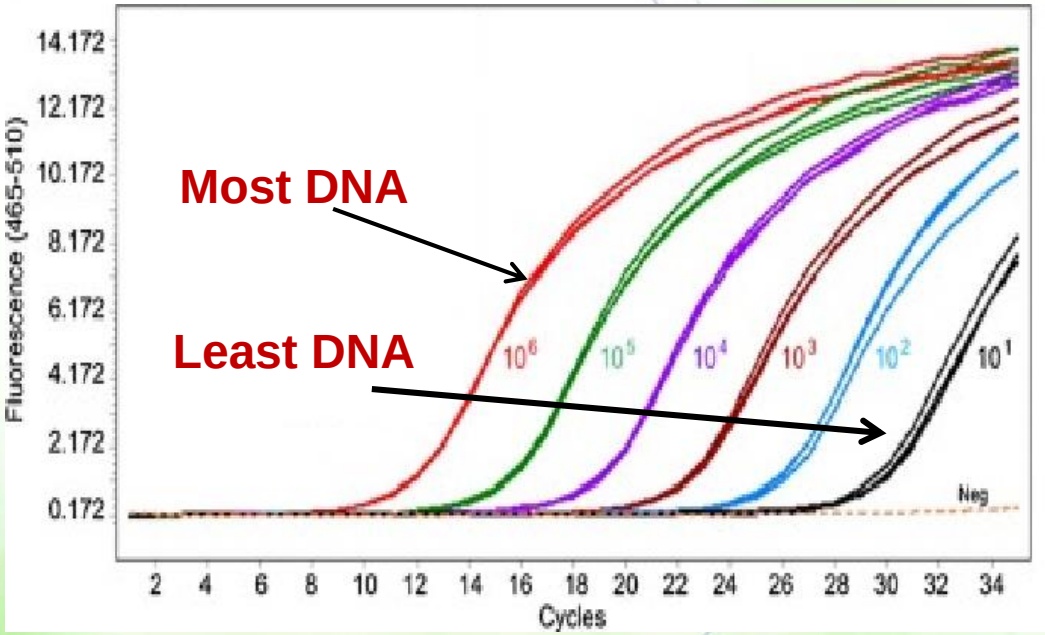
Uses of PCR

- Discovery of gene families
- Disease diagnosis
- Paternity and criminal cases. Why?
  - An individual DNA profile is highly distinctive because many genetic loci are highly variable within a population.
- Viral and bacterial load: the quantity of virus in a given volume. How?
  - Quantitative PCR

Quantitative PCR (qPCR)



- SYBR green binds to double-stranded DNA and fluoresces only when bound.
- A way of relative quantitation of amount of DNA in a sample is by amplifying it in the presence of SYBR green.
- The higher the amount of DNA, the sooner it is detected.



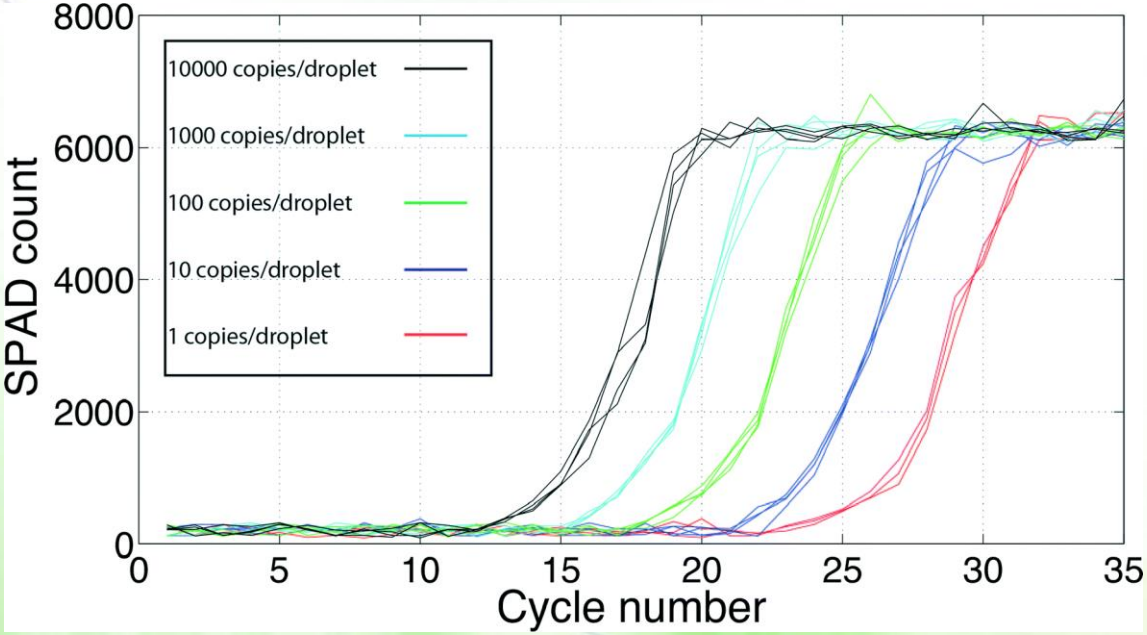
Analysis of gene expression

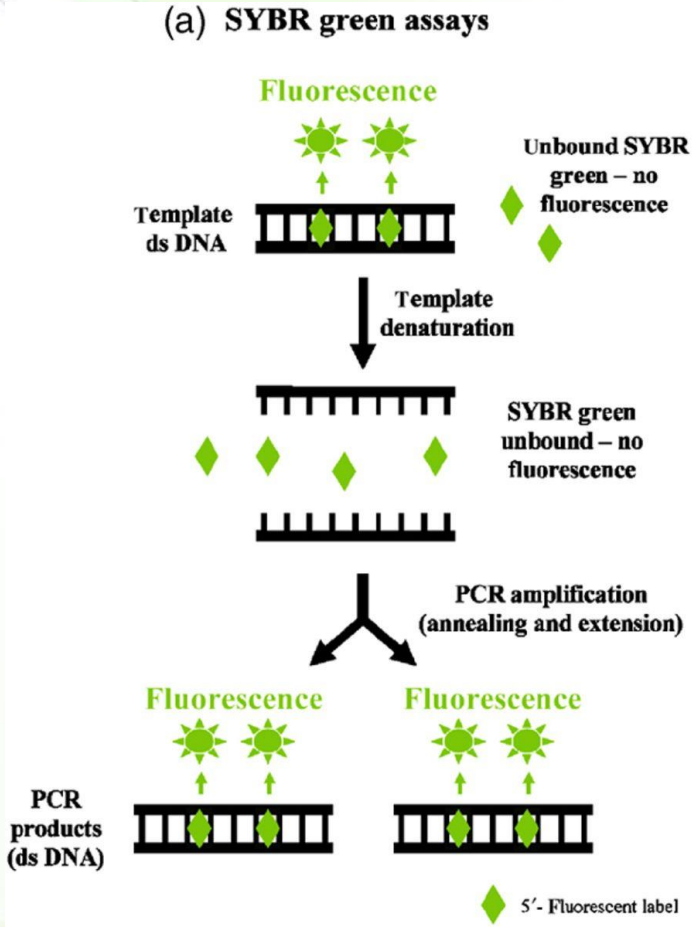
RNA level

First make a complementary DNA (cDNA)

Quantitative real-time PCR of mRNA

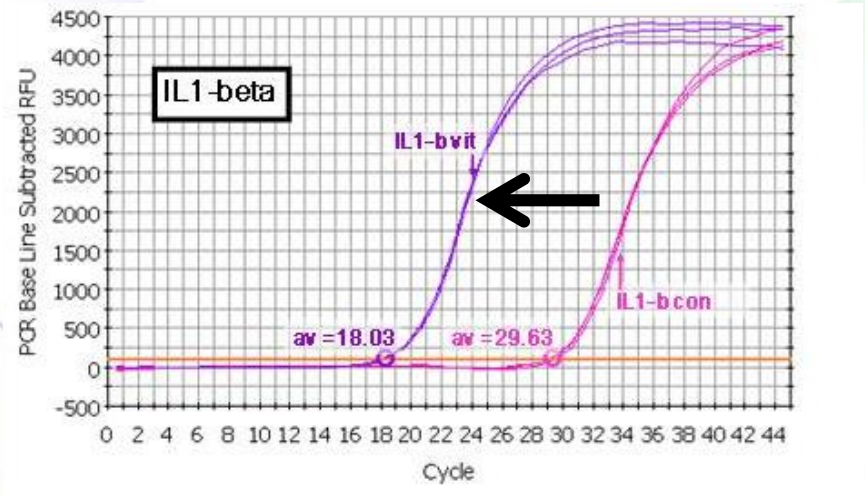
- Another way of relative quantitation of RNA expression is by converting RNA into cDNA followed by PCR in the presence of SYBR green.
- The higher the amount of RNA (cDNA), the sooner it is detected.



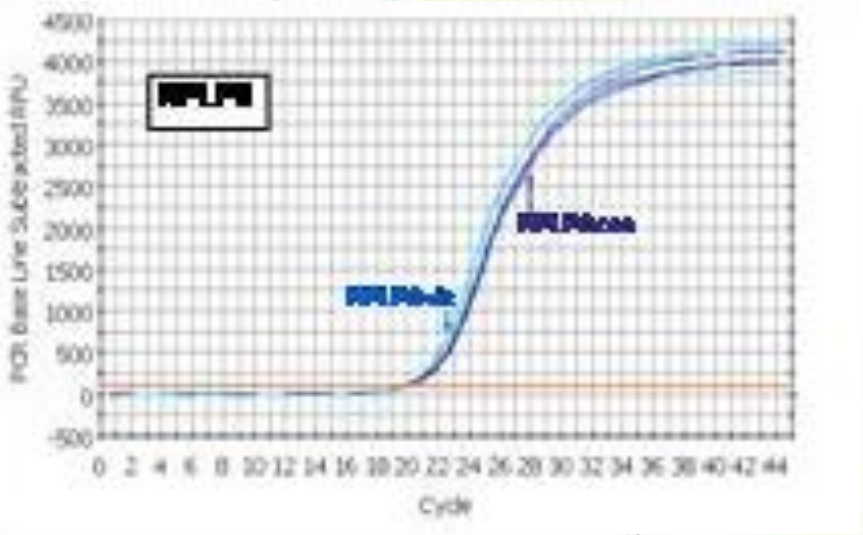


Example

A gene of interest

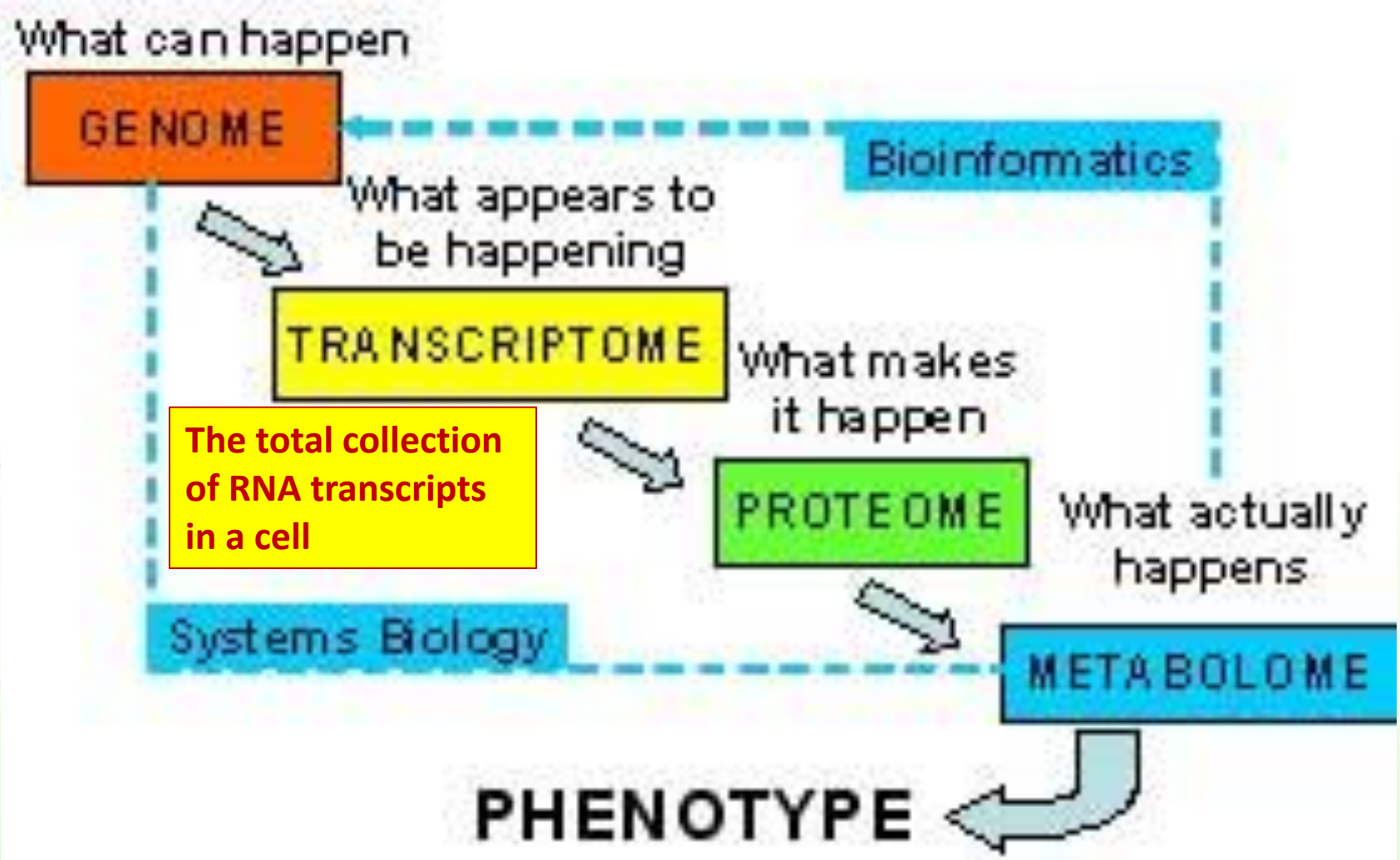


Housekeeping gene



Unaltered expression

The science of -omics



Studying the transcriptome

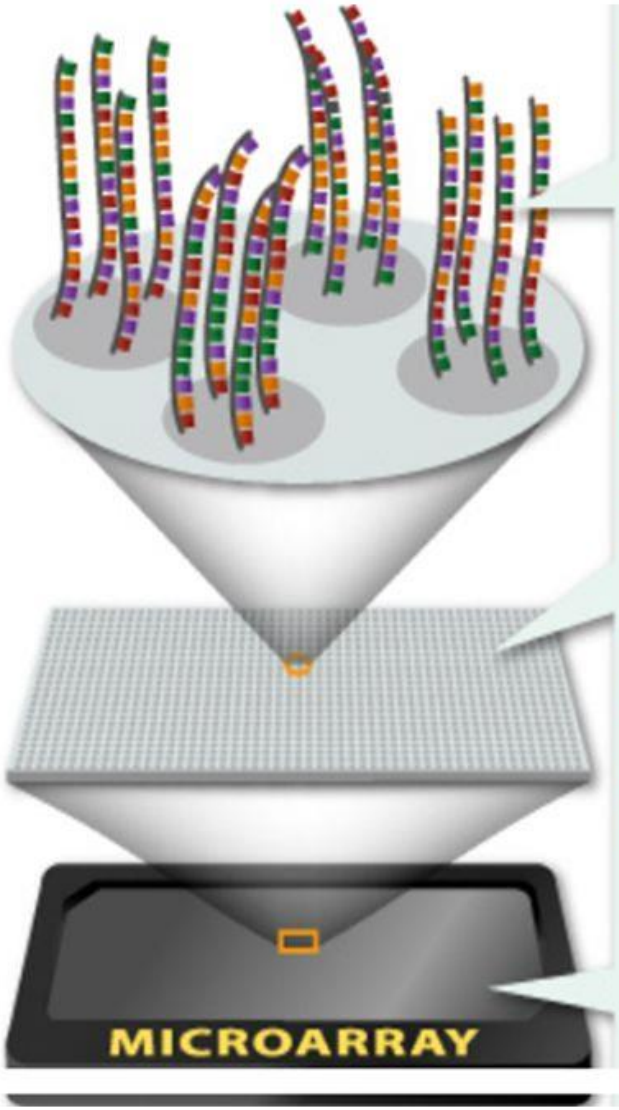
- One such method in studying transcriptomes is DNA microarrays, which allow the analysis of the RNA products of thousands of genes all at once.

- By examining the expression of so many genes simultaneously, we can understand gene expression patterns in physiological and pathological states.

DNA microarrays

- DNA microarrays are glass microscope slides spotted with up to tens of thousands of DNA fragments in an area the size of a fingernail.
- The exact sequence and position of every DNA fragment on the array is known.

- http://learn.genetics.utah.edu/content/labs/microarray/
- http://www.sumanasinc.com/webcontent/animations/content/dnachips.html



A DNA micorarray allows scientists to perform an experiment on thousands of genes at the same time.

Each spot on a microarray contains multiple identical strands of DNA.

The DNA sequence on each spot is unique.

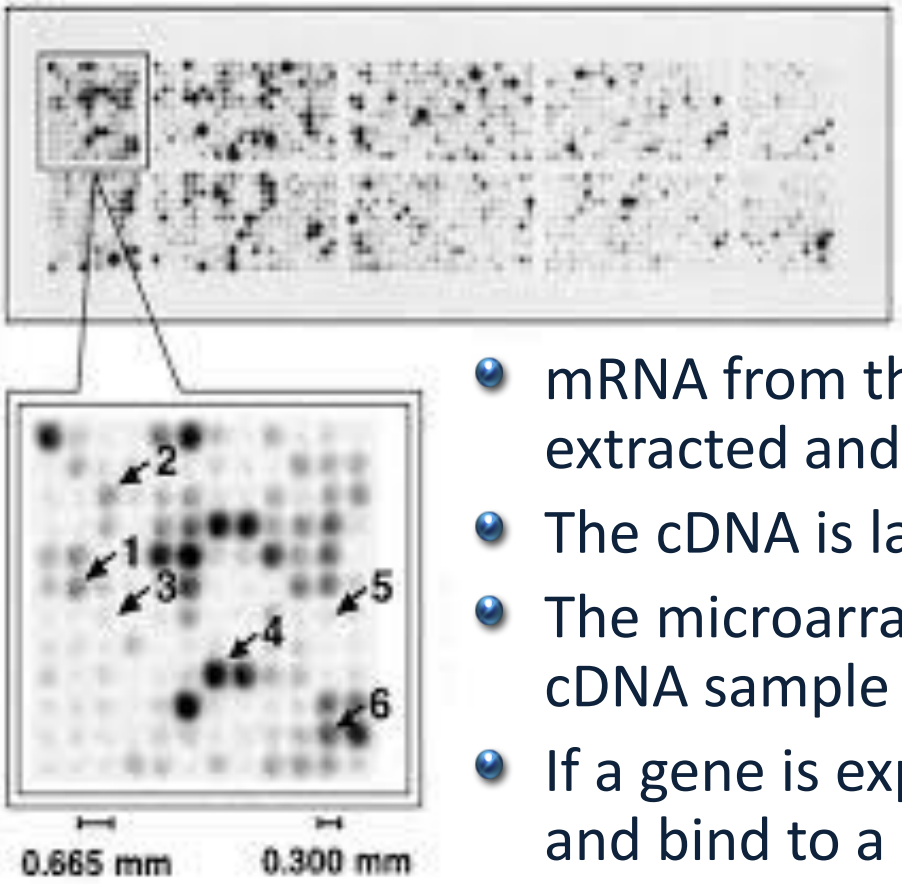
Each spot represents one gene.

Thousands of spots are arrayed in orderly rows and columns on a solid surface (usually glass).

The precise location and sequence of each spot is recorded in a computer database.

Microarrays can be the size of a microscope slide, or even smaller.

A DNA microarray

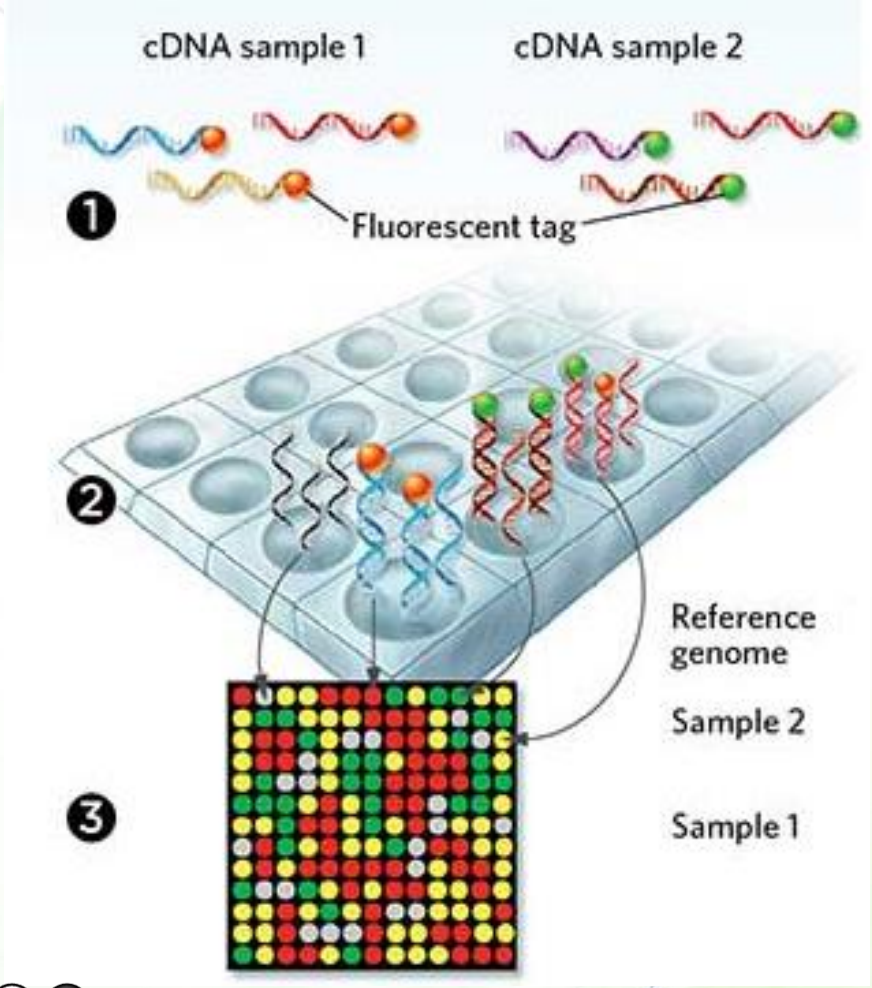


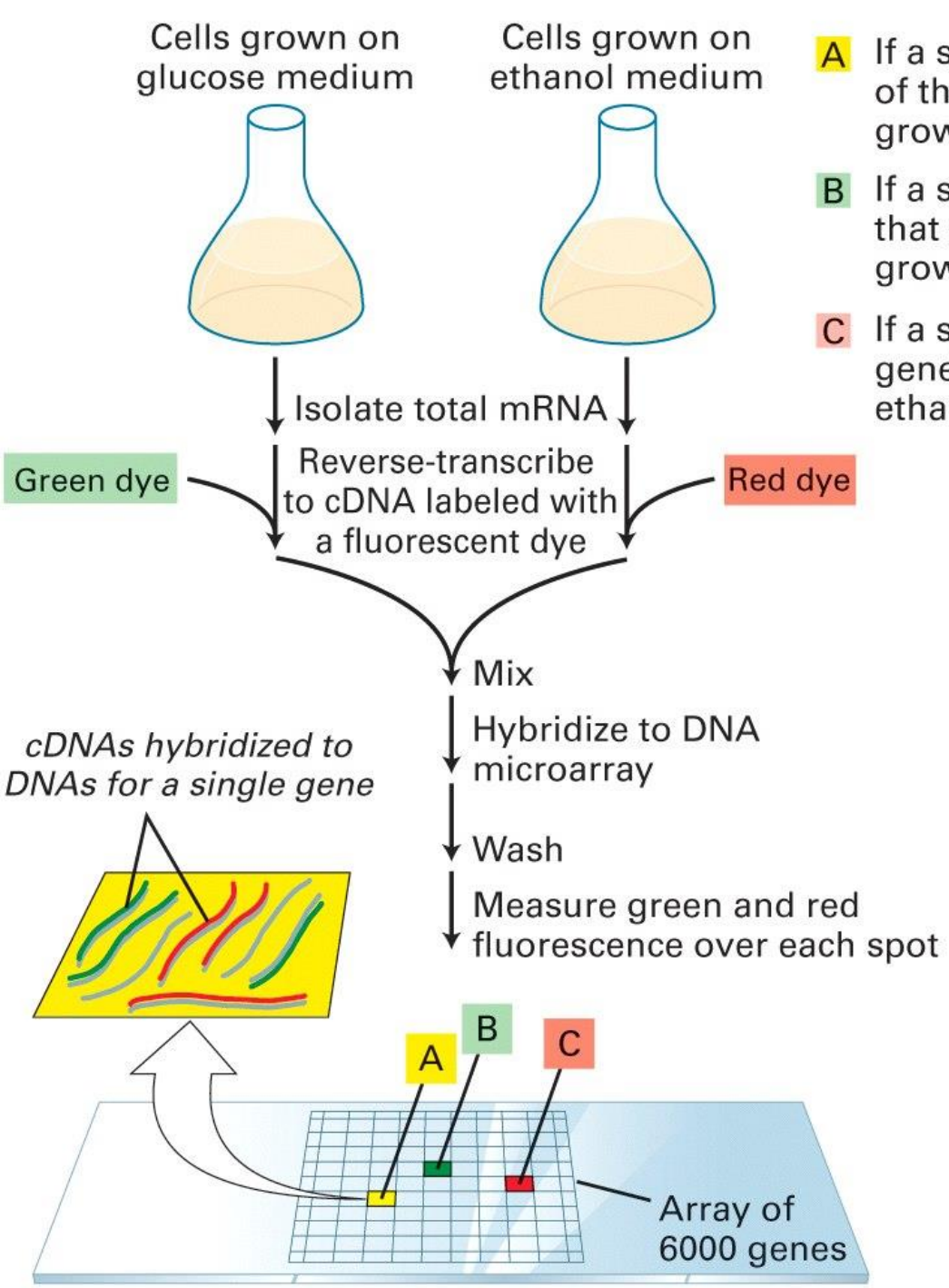
This is done for a single sample using radioactively labeled cDNA.

- mRNA from the cells being studied is first extracted and converted to cDNA.
- The cDNA is labeled with a radioactive probe.
- The microarray is incubated with the labeled cDNA sample for hybridization to occur.
- If a gene is expressed, then the cDNA will exist and bind to a specific complementary DNA fragment on the microarray.
- Binding can be detected since the cDNA is labeled and expression is determined.

Comparative expression

- In order to compare expression of genes two different samples, the cDNA molecules are fluorescently labeled with different colors (green and red) and added to the array.
- An increases in the amount of a RNA molecule in one sample versus the other is reflected by an increase the amount of produced cDNA and an increase in fluorescence in the bound spot.

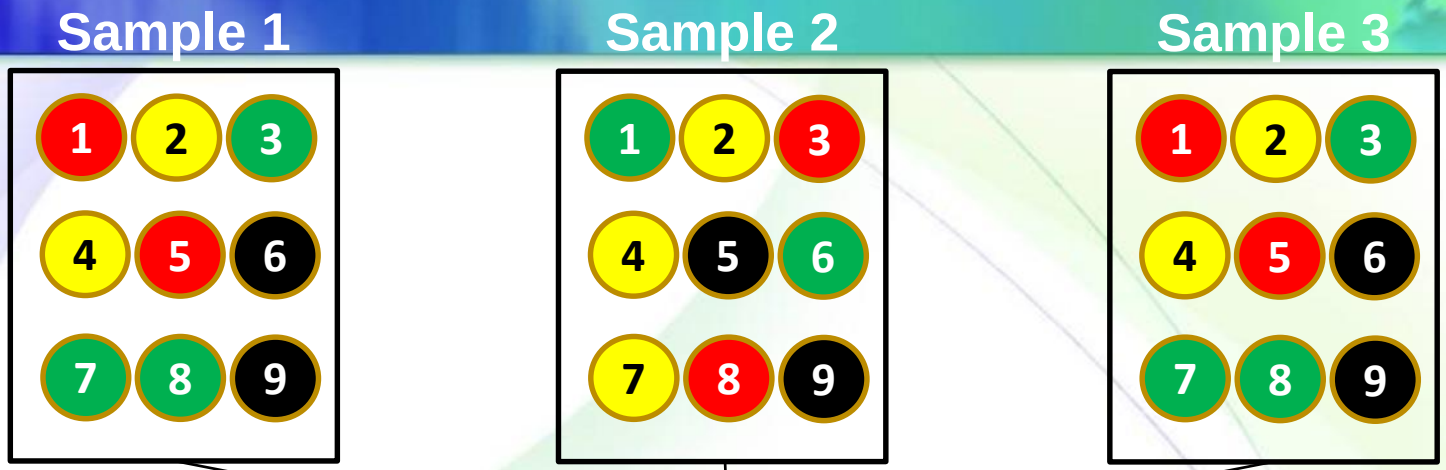




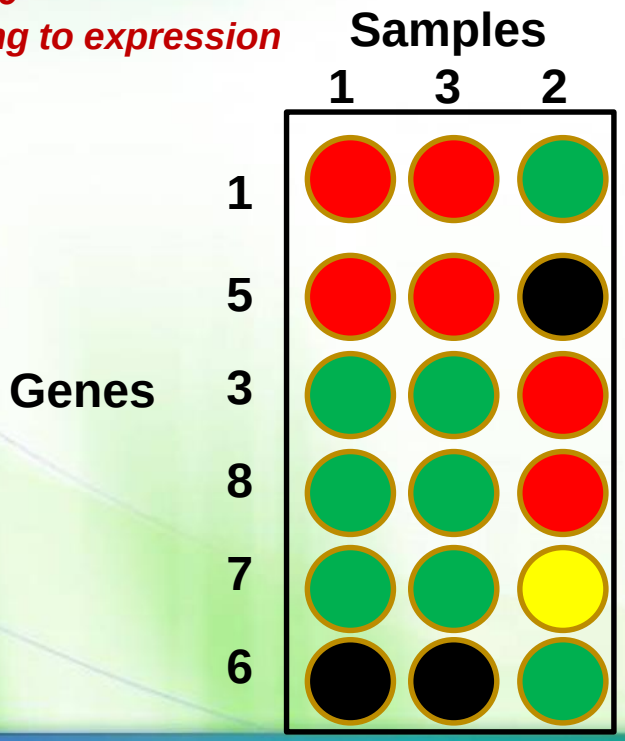
A If a spot is yellow, expression of that gene is the same in cells grown either on glucose or ethanol

B If a spot is green, expression of that gene is greater in cells grown in glucose

C If a spot is red, expression of that gene is greater in cells grown in ethanol

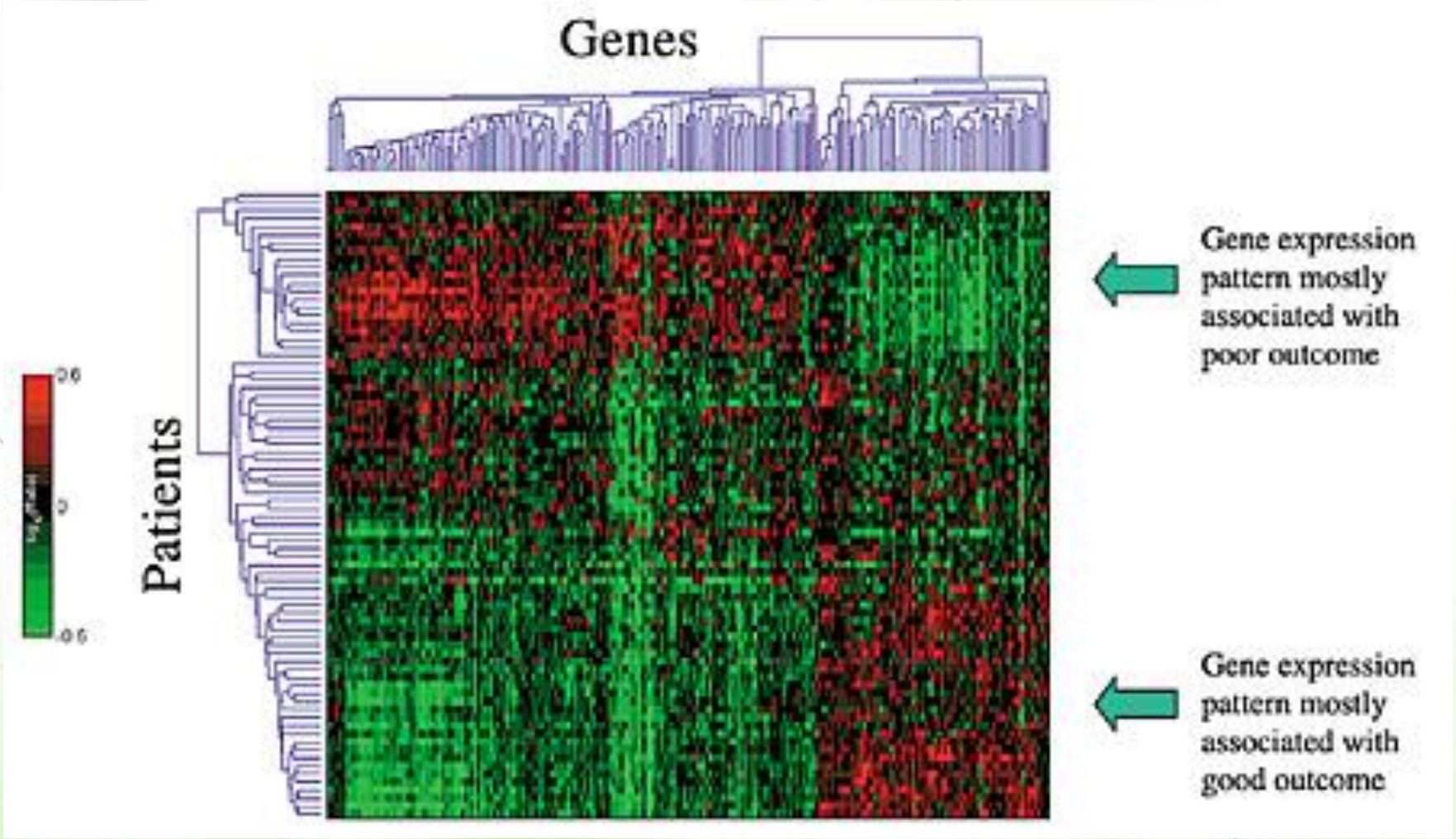


Combine results
Eliminate samples 2, 4,& 9
Cluster samples according to expression



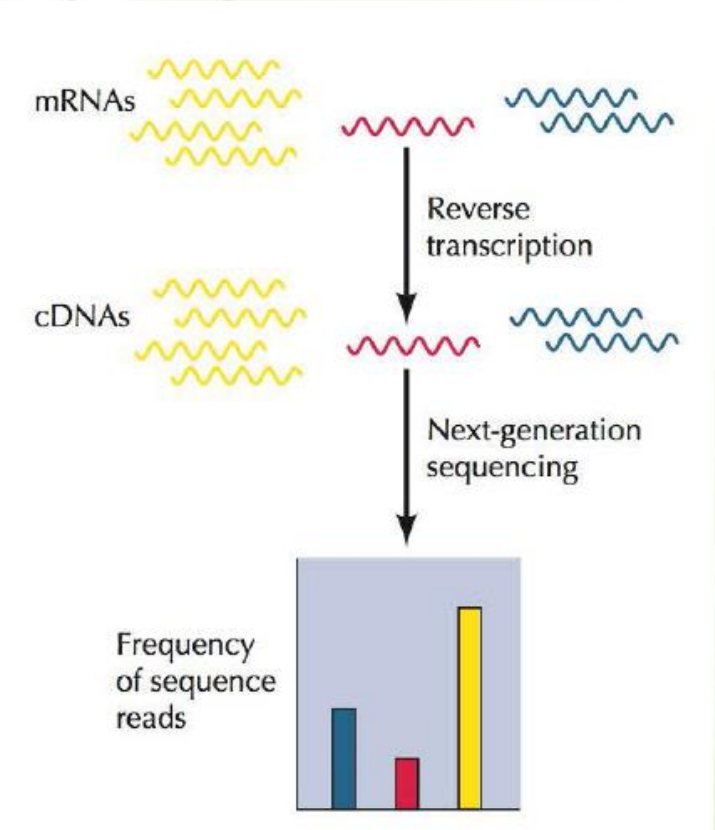
This can be done at a larger scale whereby samples are compared to the same control sample and a computer program combines all data illustrating differences in expression among the samples and classifying them into different groups.

DNA microarrays and breast cancer



RNA sequencing (RNA-seq)

- Cellular RNA is reverse transcribed to cDNAs, which are subjected to next-generation sequencing.
- The relative amount of each cDNA (mRNA) is indicated by the frequency at which its sequence is represented in the total number of sequnces read.



RNA-seq vs. microarray

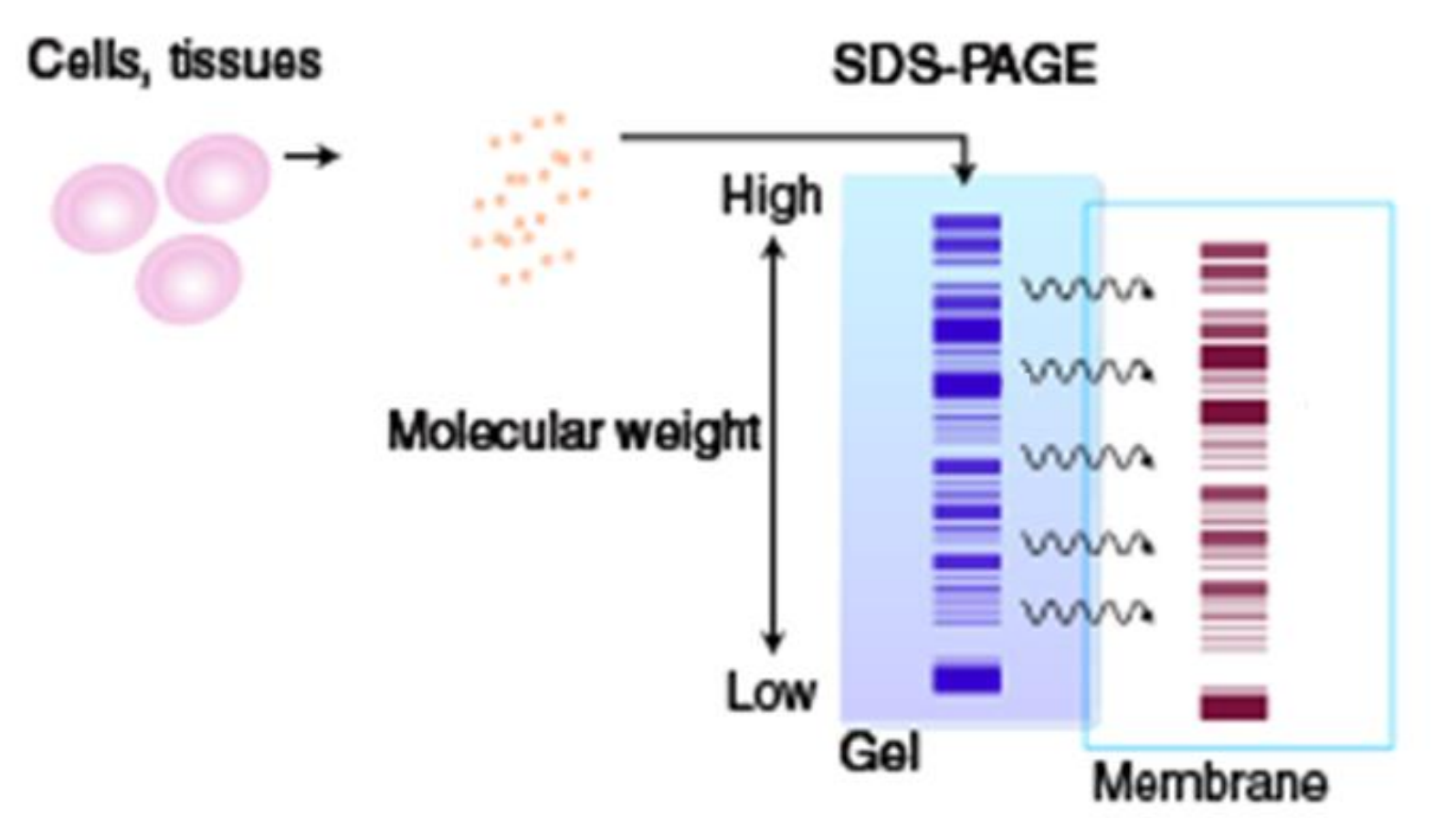
- RNA-seq can be used to
  - characterize novel transcripts
  - Identify splicing variants
  - profile the expression levels of known transcripts
- Microarrays are limited to detect transcripts corresponding to known genomic sequences.

Analysis of gene expression

Protein level

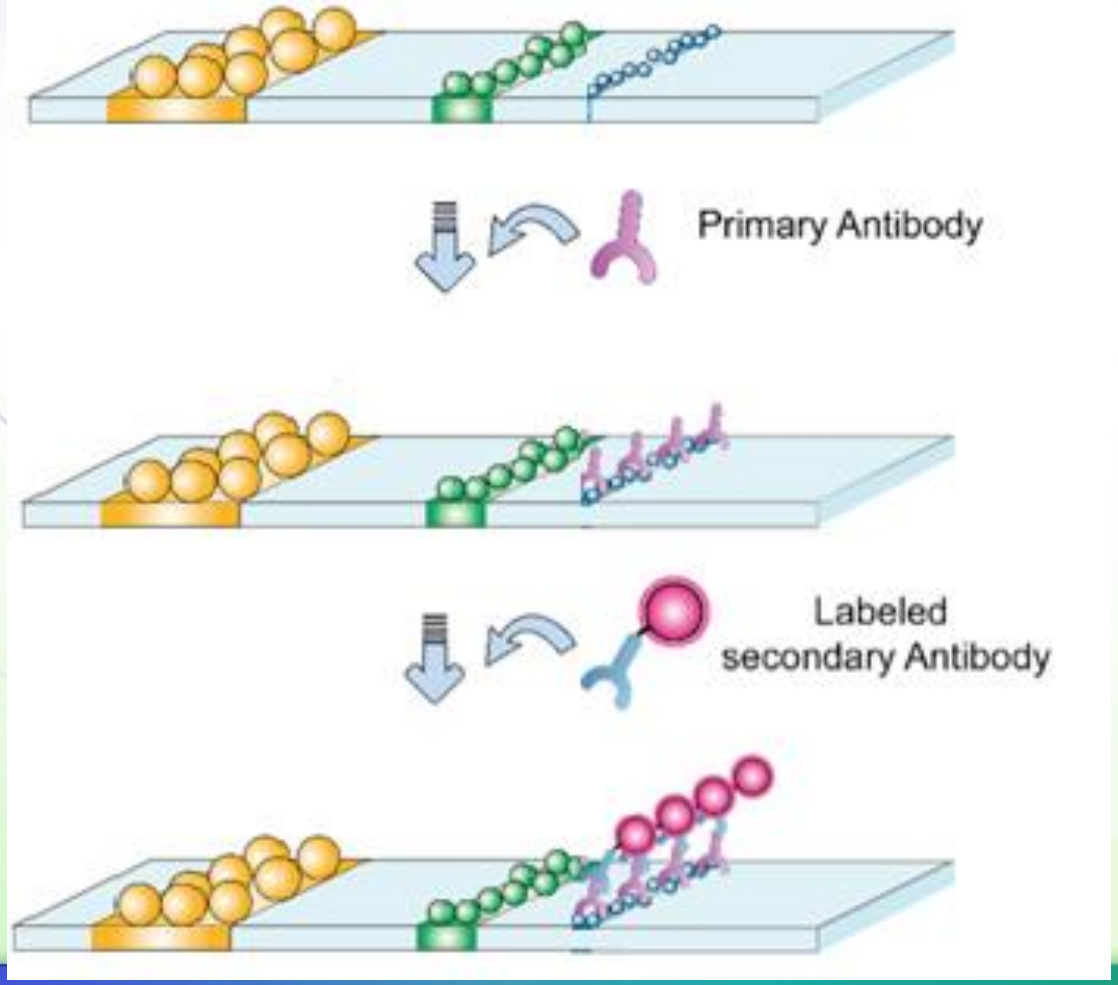
Immunoblotting (western blotting)

Step 1: Proteins are collected, separated by SDS-PAGE, and transferred to a membrane



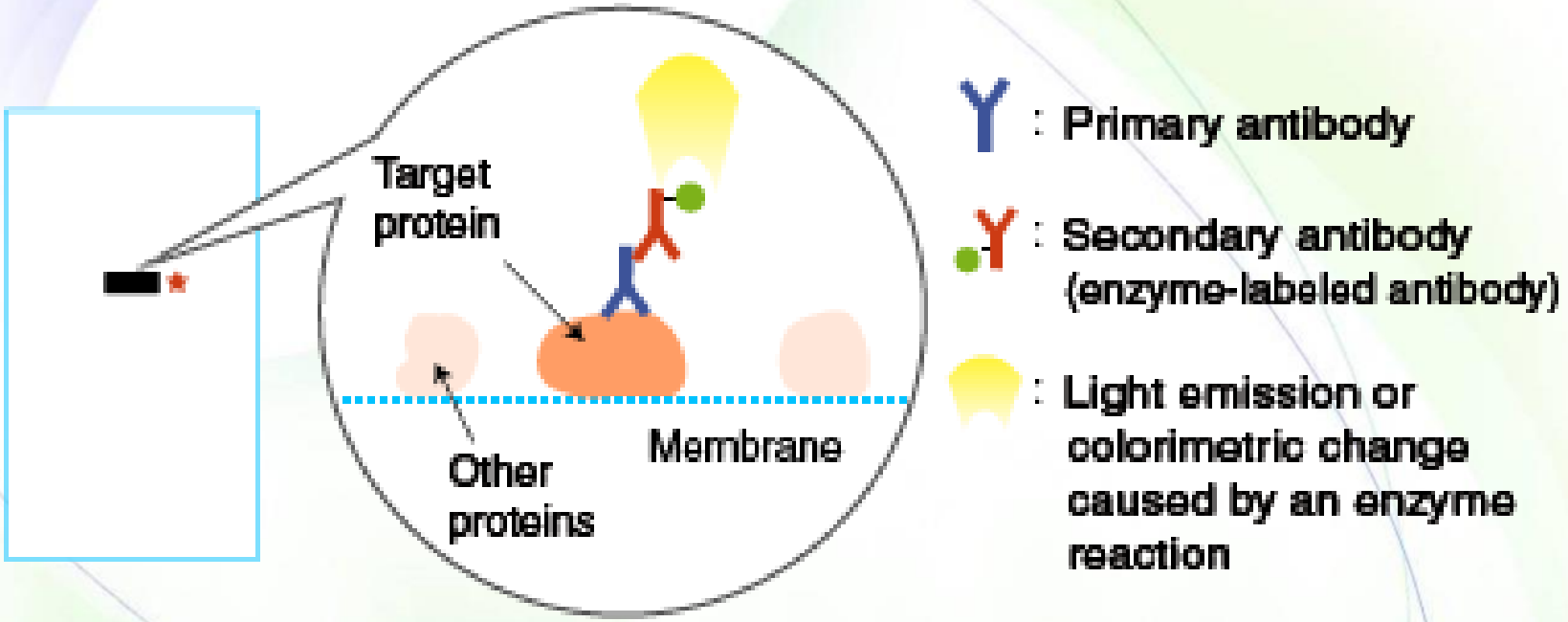
Then,

Step 2: A primary antibody that binds to the target protein is added, followed by a secondary antibody that detects the primary antibody is added.



The secondary antibody generates a signal.

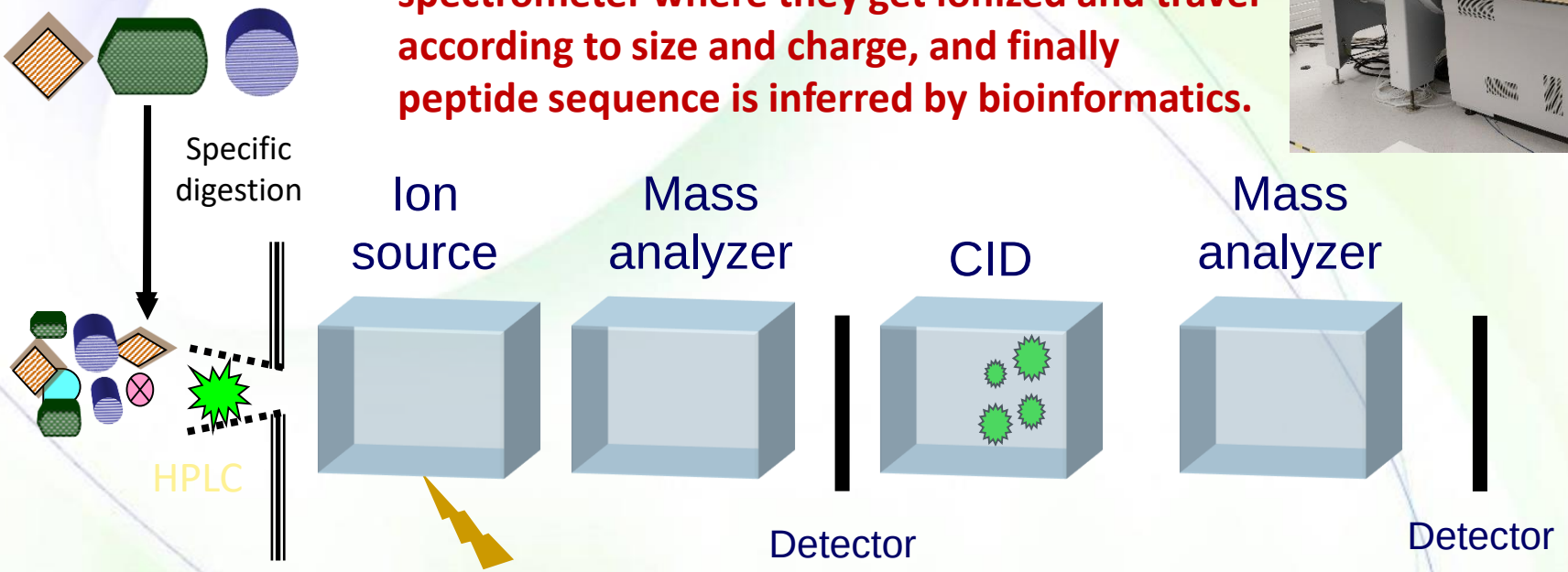
Finally, the signal is detected

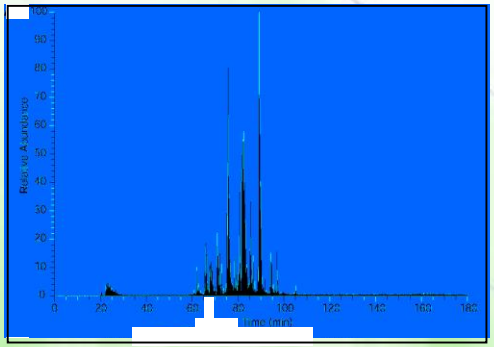


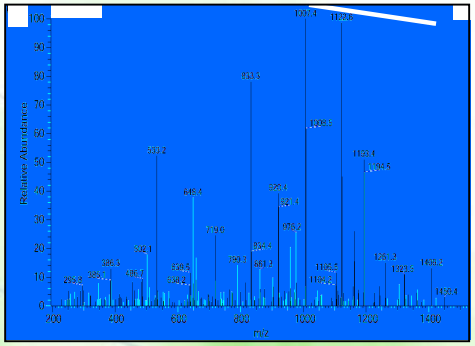
Thinks: what information do we get?

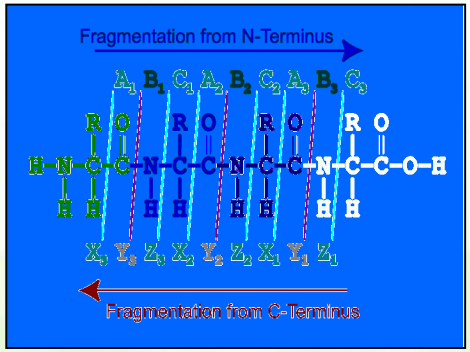
Mass spectrometry

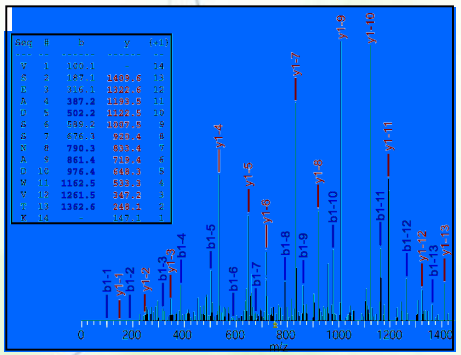
Proteins are digested, separated by chromatography and injected into the mass spectrometer where they get ionized and travel according to size and charge, and finally peptide sequence is inferred by bioinformatics.



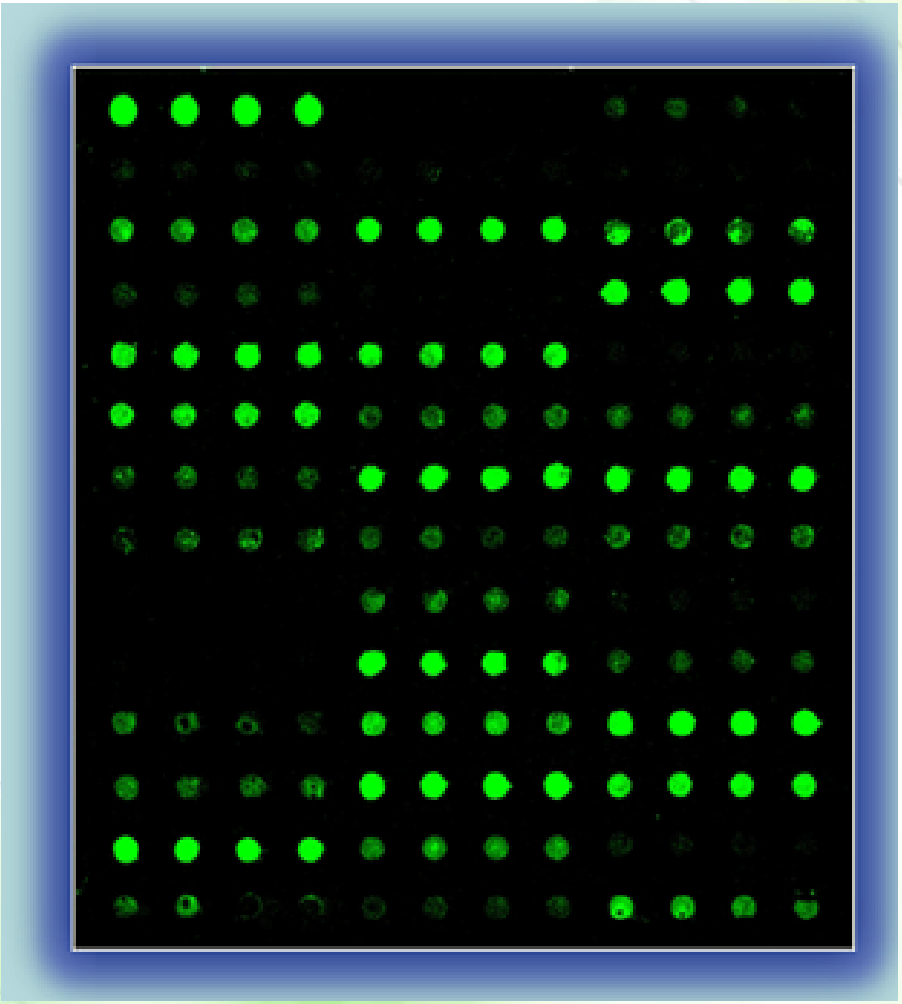






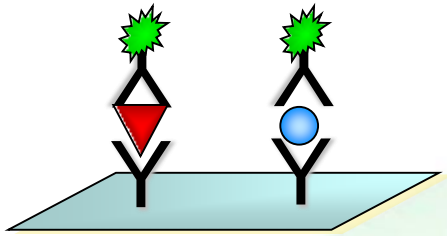


Protein arrays

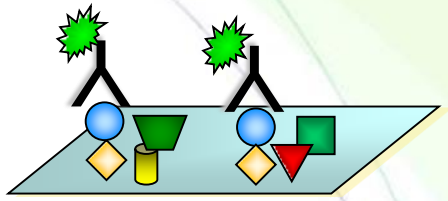


Types of protein arrays

Expression arrays

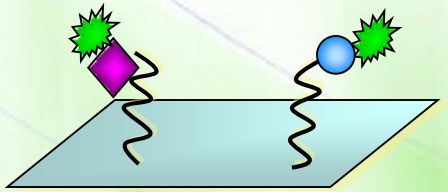


Forward-phase microarray

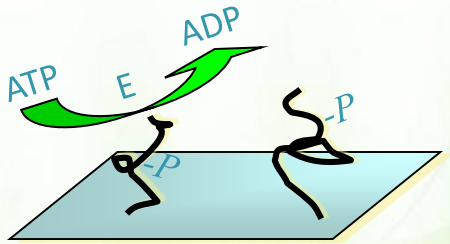


Reverse-phase microarray

Functional microarrays



Interaction microarray



Enzymatic microarray