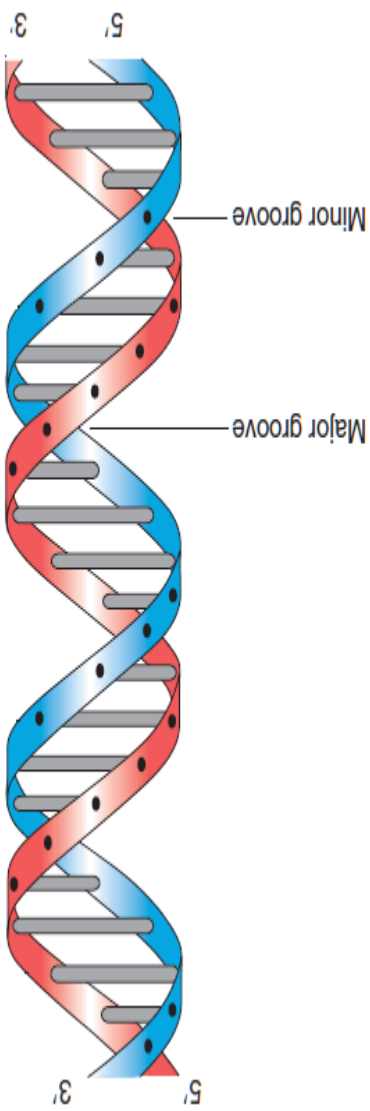
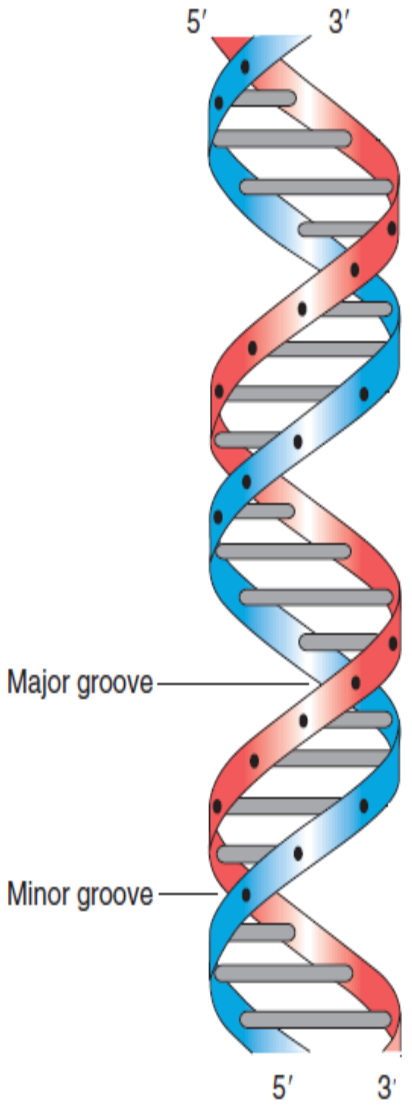
DNA Sequencing



Dr. Serageldeen A. A. Sultan
PhD in Molecular virology
Yamaguchi University, Japan (2010)
Lecturer of virology
Dept. of Microbiology
SVU, Qena, Egypt
seaas@lycos.com



What is DNA sequencing?

-Determining the precise order of nucleotides in a piece of DNA

-DNA sequence is useful in studying fundamental biological processes and in applied fields such as diagnostic or forensic research

-DNA sequencing methods have been around for 40 years, and since the mid-1970s

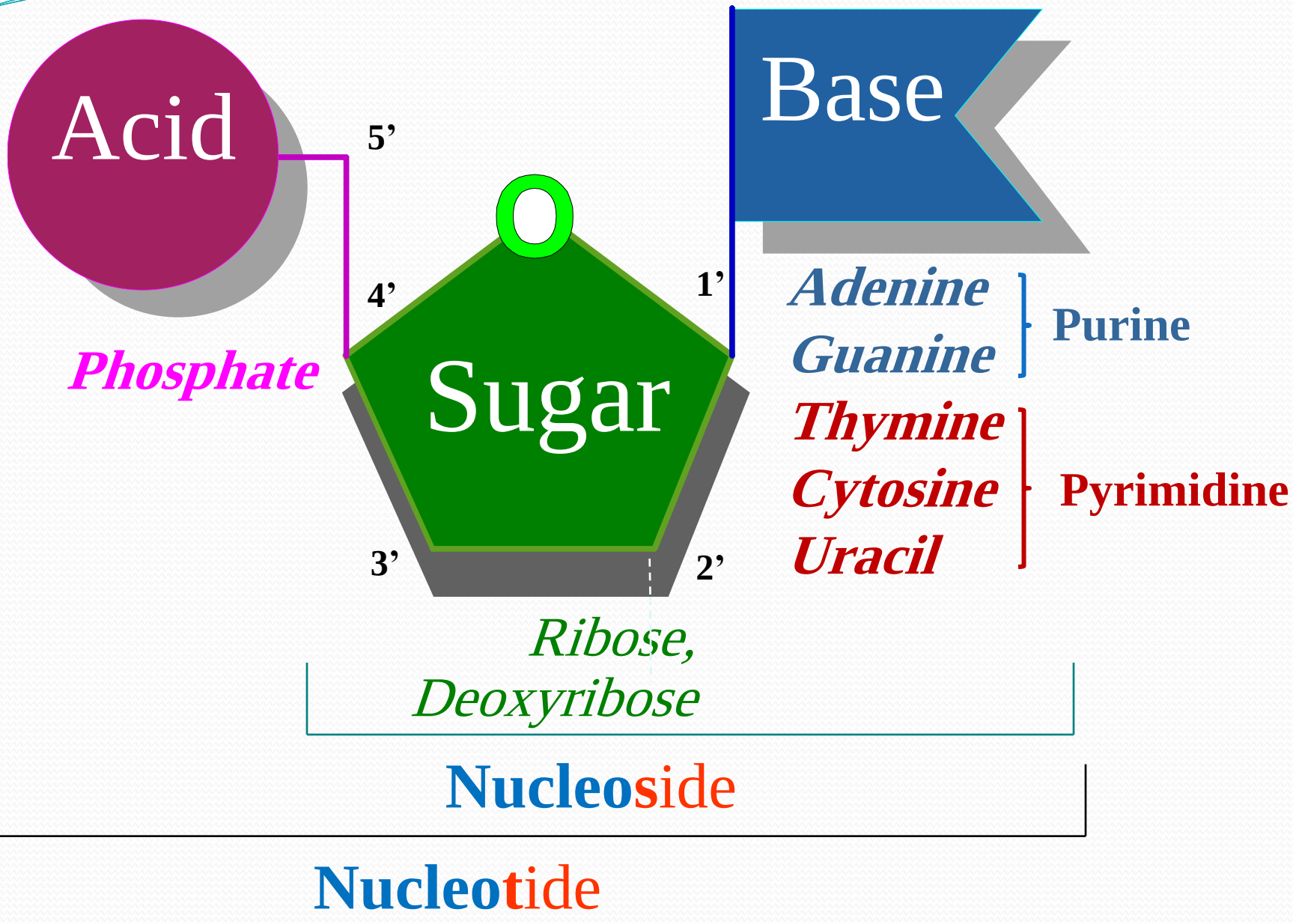
Founders of sequencing technology

Sanger

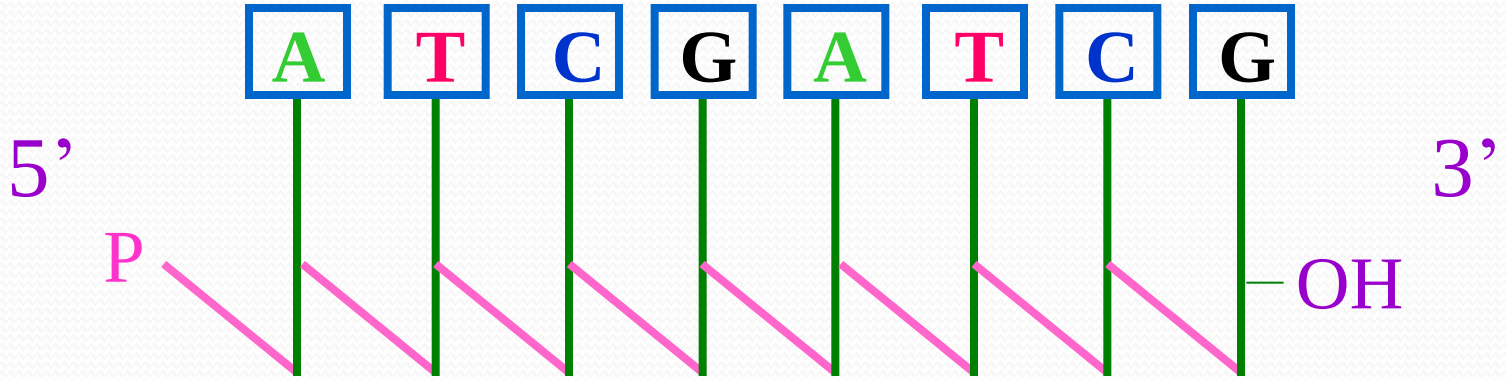
Gilbert

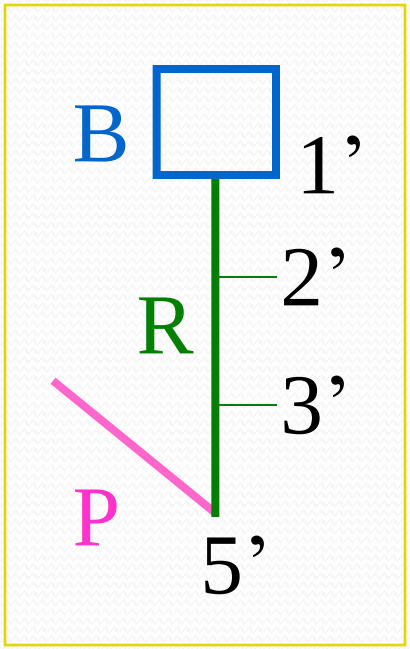
Nobel Prize in Chemistry in 1980

Basic structure of nucleic acid



Nucleotides linked by phosphodiester





5’ pApTpCpGpApTpCpG-OH 3’

5’ pATCGATCG-OH 3’

ATCGATCG

Two strands are antiparallel

5’ pApTpCpGpApTpCpG-OH 3’

3’ HO-TdAdGdCdTdAdGdC 5’

Sequencing methods

Two basic methods for DNA sequencing :-

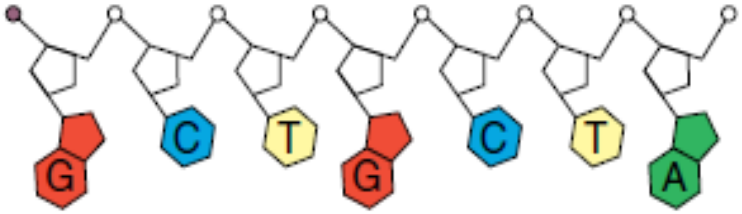
A- Chemical cleavage method (Maxam and Gilbert, 1977)

- Base-specific cleavage of DNA by certain chemicals
- Four different chemicals, one for each base
- A set of DNA fragments of different sizes
- DNA fragments contain up to 500 nucleotides

B- Enzymatic method (Sanger, 1981)

The chain cleavage reaction

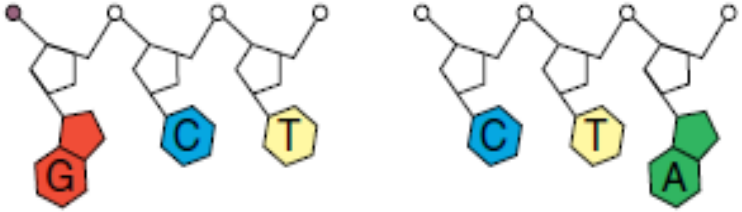
DNA labeled at one end with ^{32}p



Base modification

Release or displacement of reacted bases

Strand scission



- Dimethyl sulfate (DMS) mehylates G -Acid (A)

-Hydrazine (C)
- Hydrazine & NaCl (T)
- Piperidine

The fragments created by chain cleavage at guanines

^{32}pGpCpTpGpCpTpApGpGpTpGpCpCpGpApGpC

G G G G G G

^{32}p

^{32}pGpCpTp

^{32}pGpCpTpGpCpTpAp

^{32}pGpCpTpGpCpTpApGp

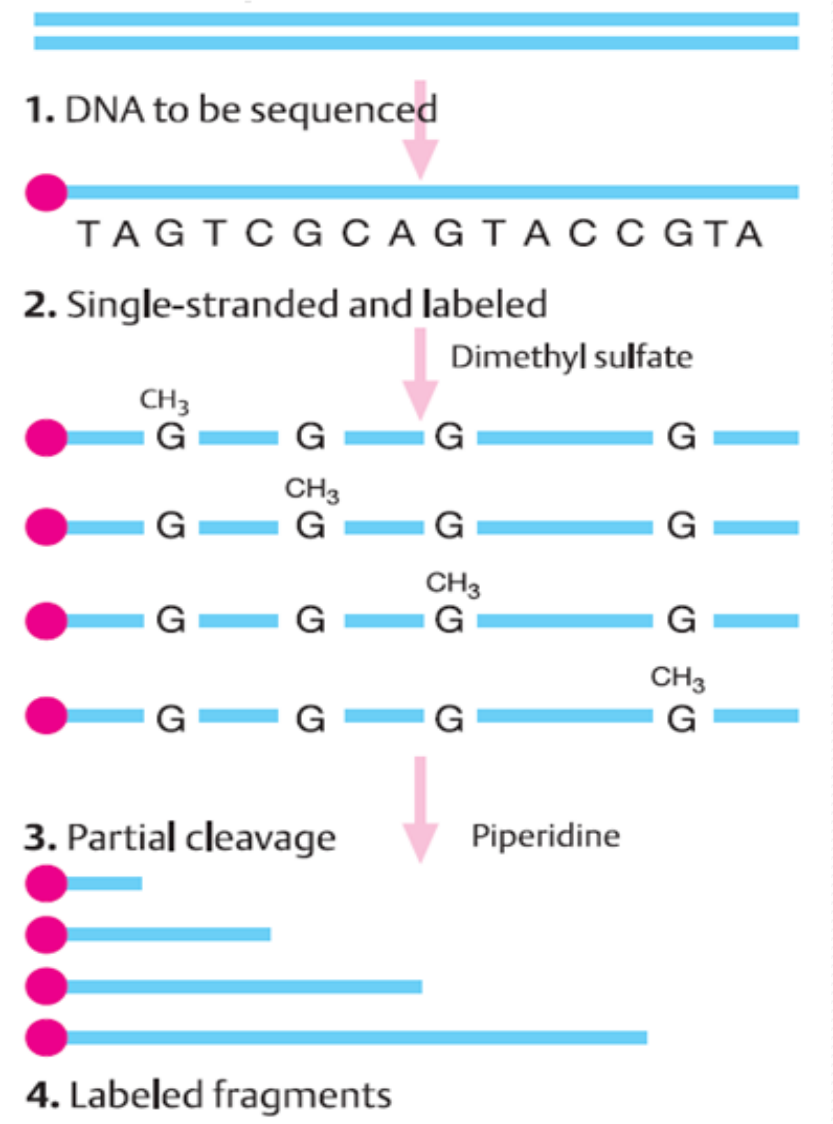
^{32}pGpCpTpGpCpTpApGpGpTp

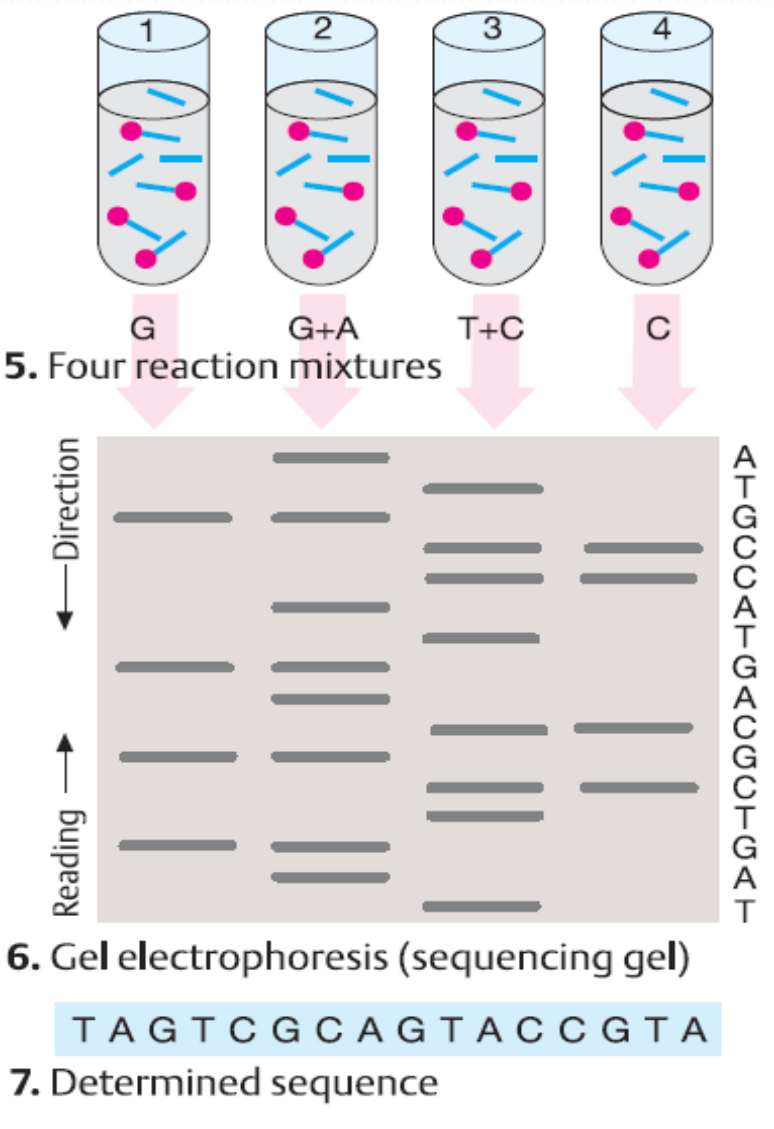
^{32}pGpCpTpGpCpTpApGpGpTpGpCpCp

^{32}pGpCpTpGpCpTpApGpGpTpGpCpCpGpAp

^{32}pGpCpTpGpCpTpApGpGpTpGpCpCpGpApGpC

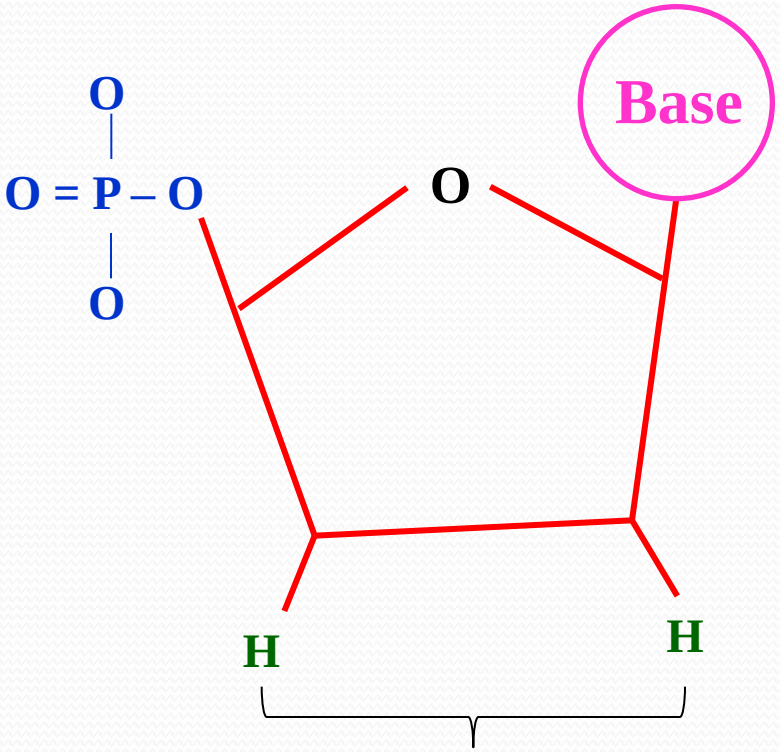
Chemical degradation method (Maxam–Gilbert method)





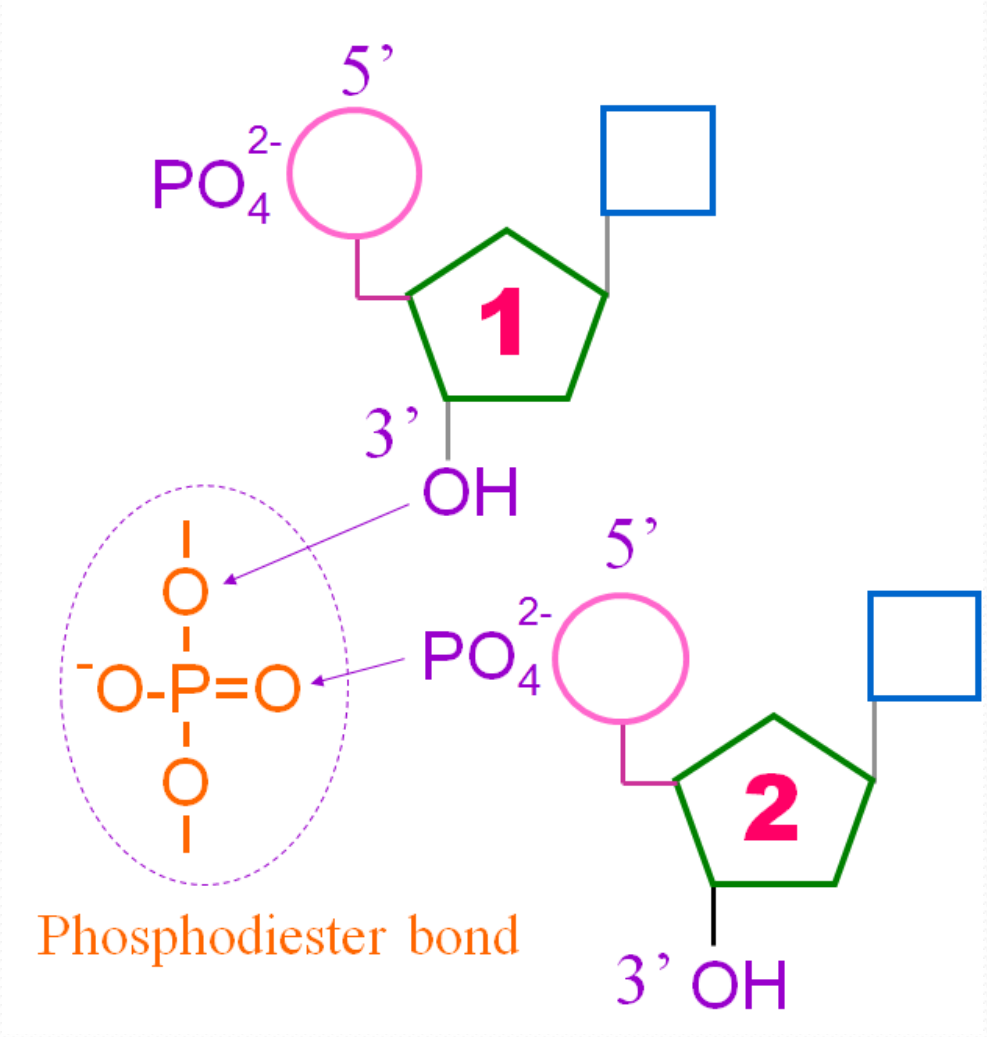
The Sanger DNA sequencing method

- Uses dideoxy nucleotides to terminate DNA synthesis.
- DNA synthesis reactions in four separate tubes
- Radioactive dATP is also included in all the tubes so the DNA products will be radioactive.
- Yielding a series of DNA fragments whose sizes can be measured by electrophoresis.
- The last base in each of these fragments is known.

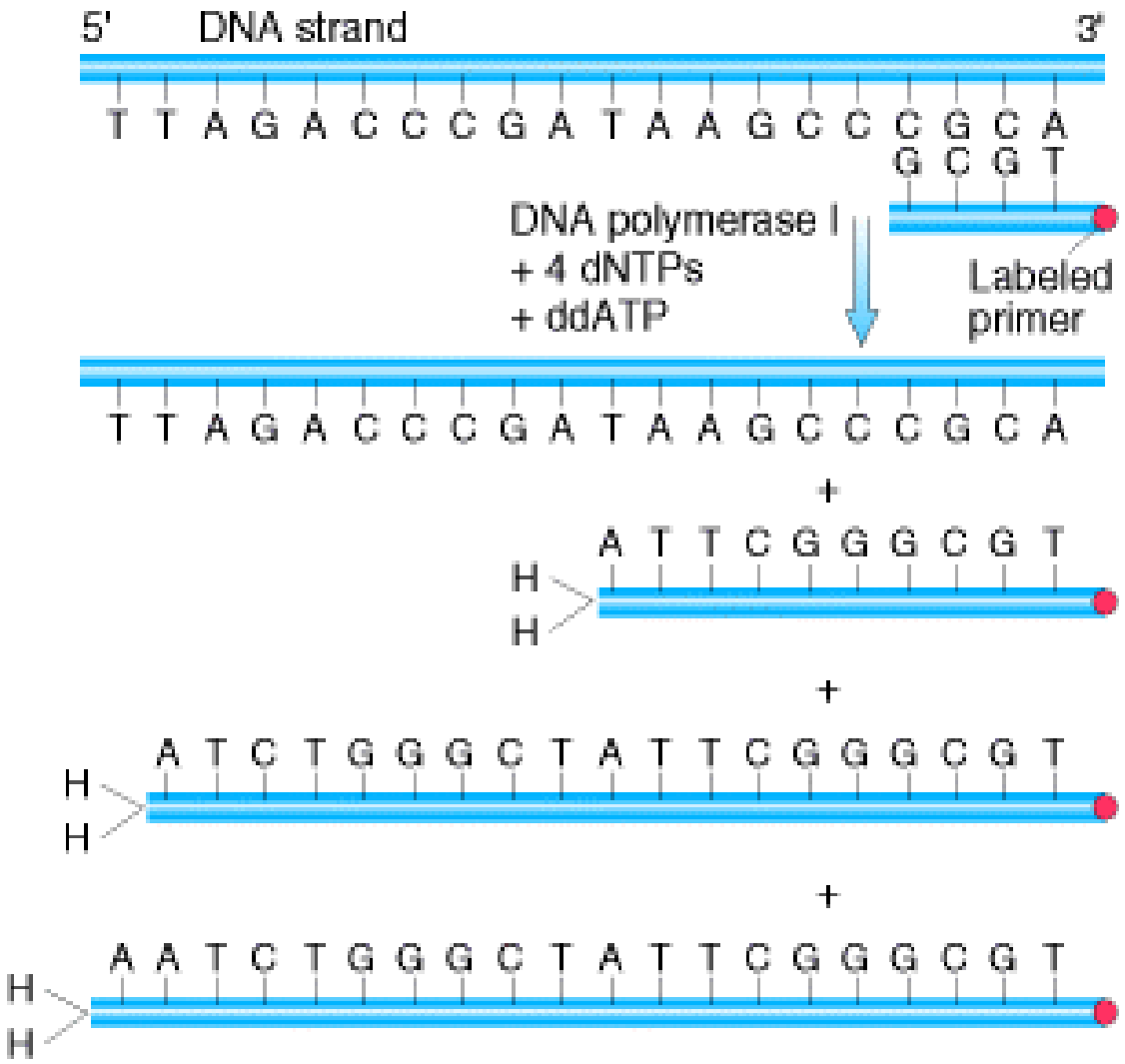


2’, 3’ dideoxy nucleotide

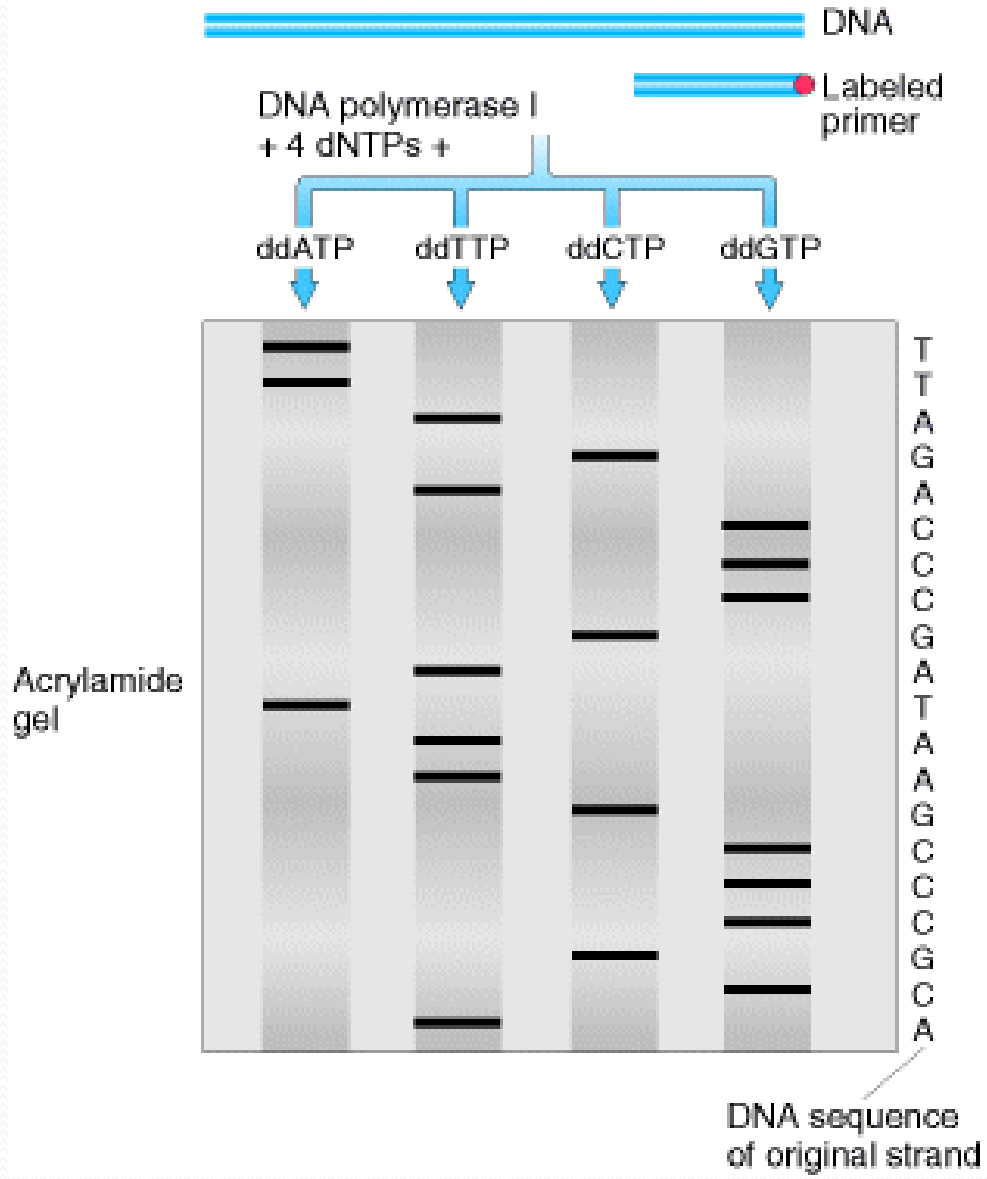
Can not form phosphodiester bound with next coming dNTP



The dideoxy sequencing method (Sanger method)



A labeled primer is used to initiate DNA synthesis. The addition of four different dideoxy nucleotides randomly arrests synthesis.



The resulting fragments are separated electrophoretically and subjected to autoradiography

Recent methods of chain termination sequencing

- **Thermal cycler sequencing**
- **Automated DNA sequencing**
- **Pyrosequencing**
- **Sequencing by hybridization**

Automated DNA sequencing

-The primer extension reactions are run in the same way as in the manual method

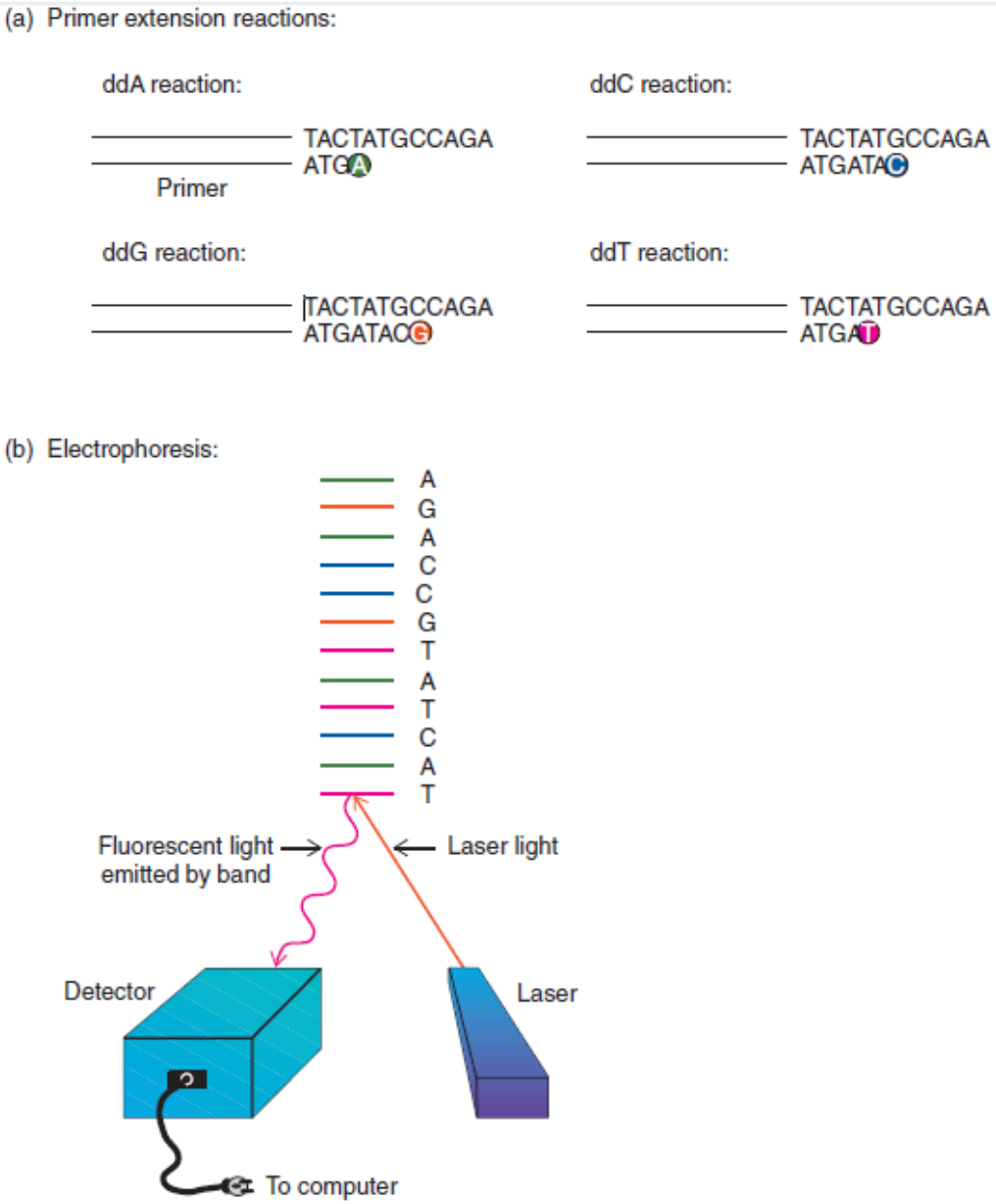
-Reaction carried out in one tube and all possible products are actually produced

- The various reaction products separate according to size on gel electrophoresis

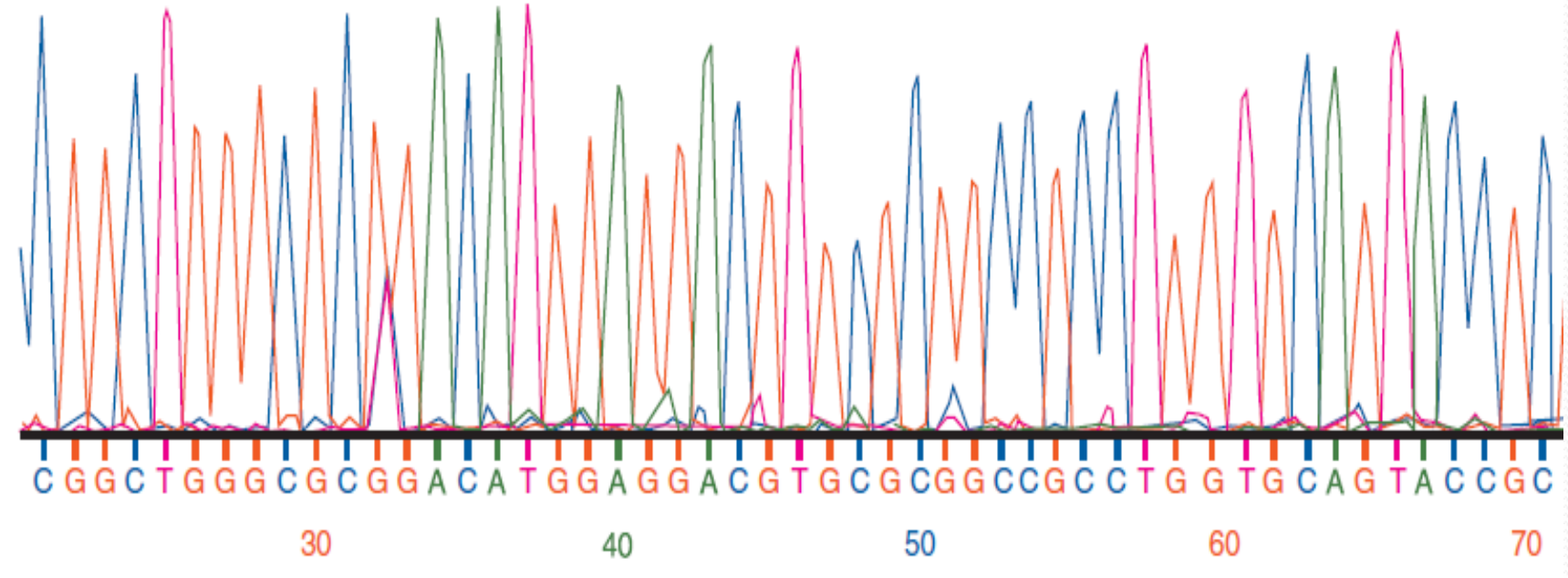
-The bands are color-coded according to the termination reaction that produced them

-A laser scanner excites the fluorescent tag on each band as it passes by, and a detector analyzes the color of the resulting emitted light

- Each colored peak is a plot of the fluorescence intensity of a band as it passes through the laser beam

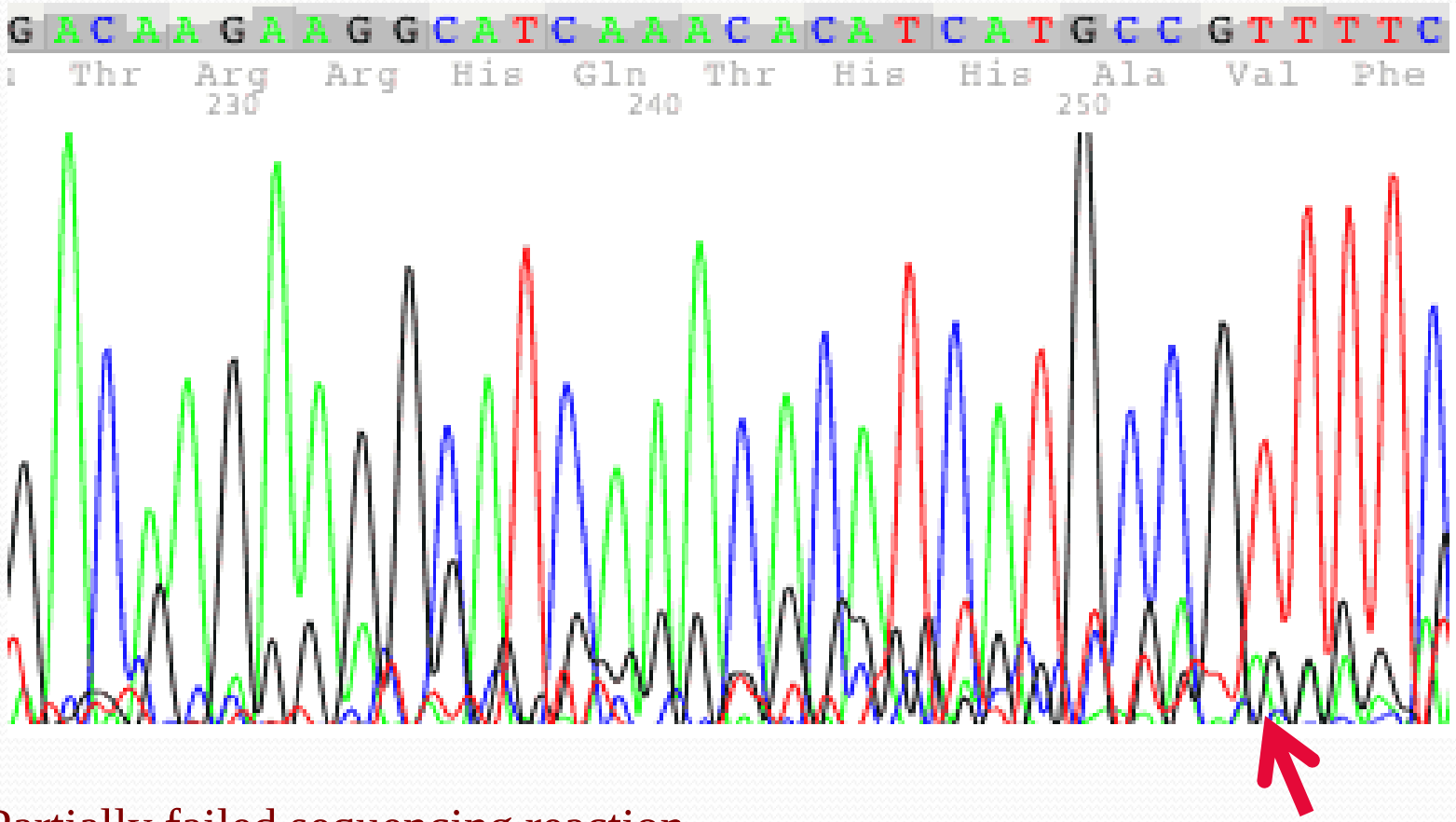

(a) Primer extension reactions:
ddA reaction:
TACTATGCCAGA
ATGA
Primer
ddC reaction:
TACTATGCCAGA
ATGATAC
ddG reaction:
TACTATGCCAGA
ATGATACG
ddT reaction:
TACTATGCCAGA
ATGAT
(b) Electrophoresis:
A
G
A
C
C
G
T
A
T
C
A
T
Fluorescent light emitted by band
Laser light
Detector
Laser
To computer

Printout of an automated DNA sequencing



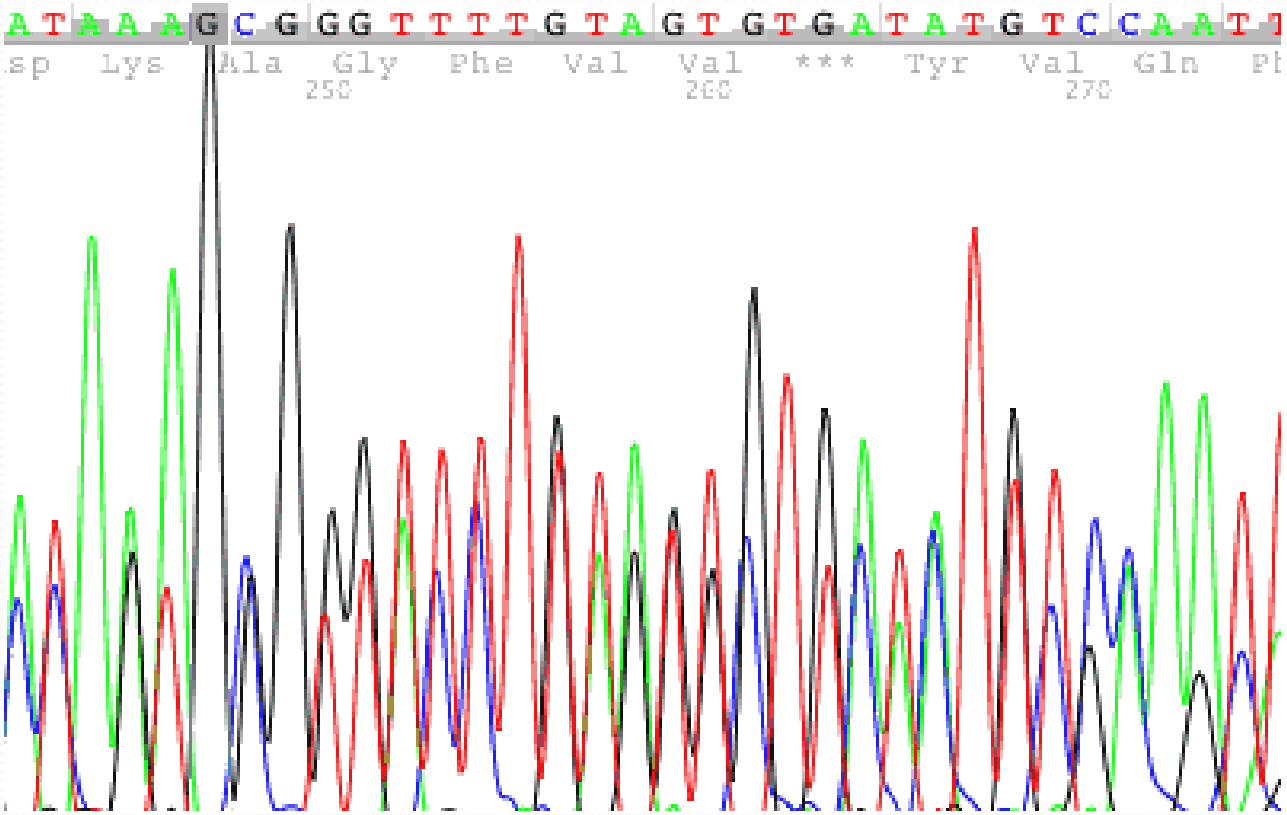
How can you evaluate sequencing results?

Noisy or weak signal



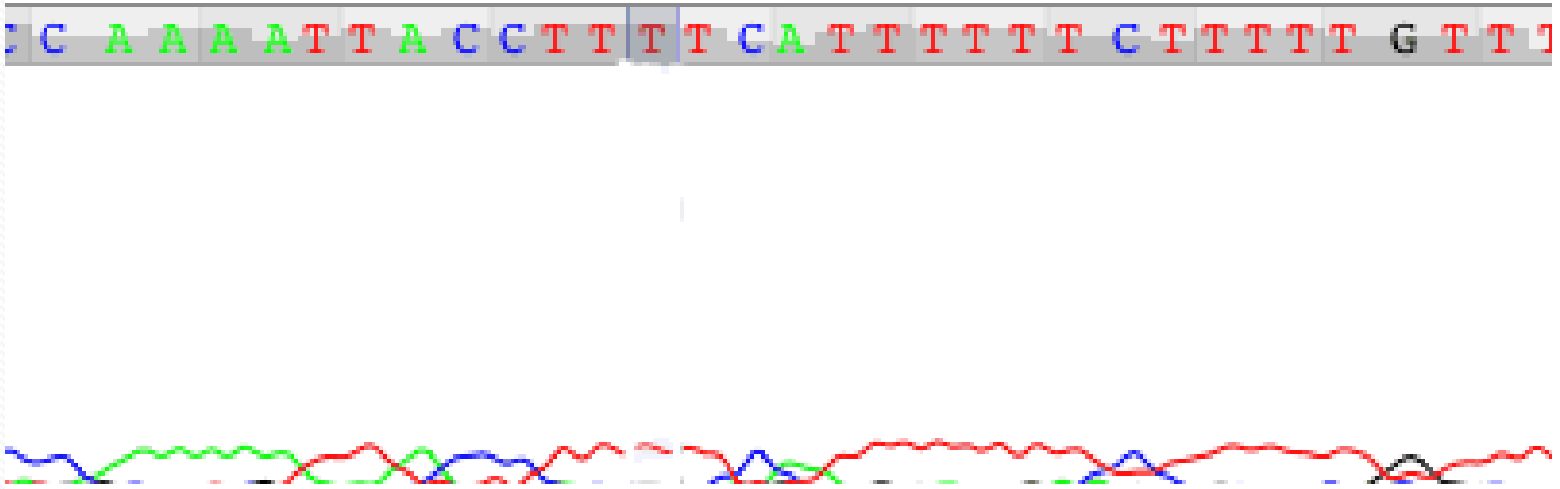
- Partially failed sequencing reaction.
- Too much or too little DNA.
- Partial loss of sample during clean-up

Mixed DNA templates



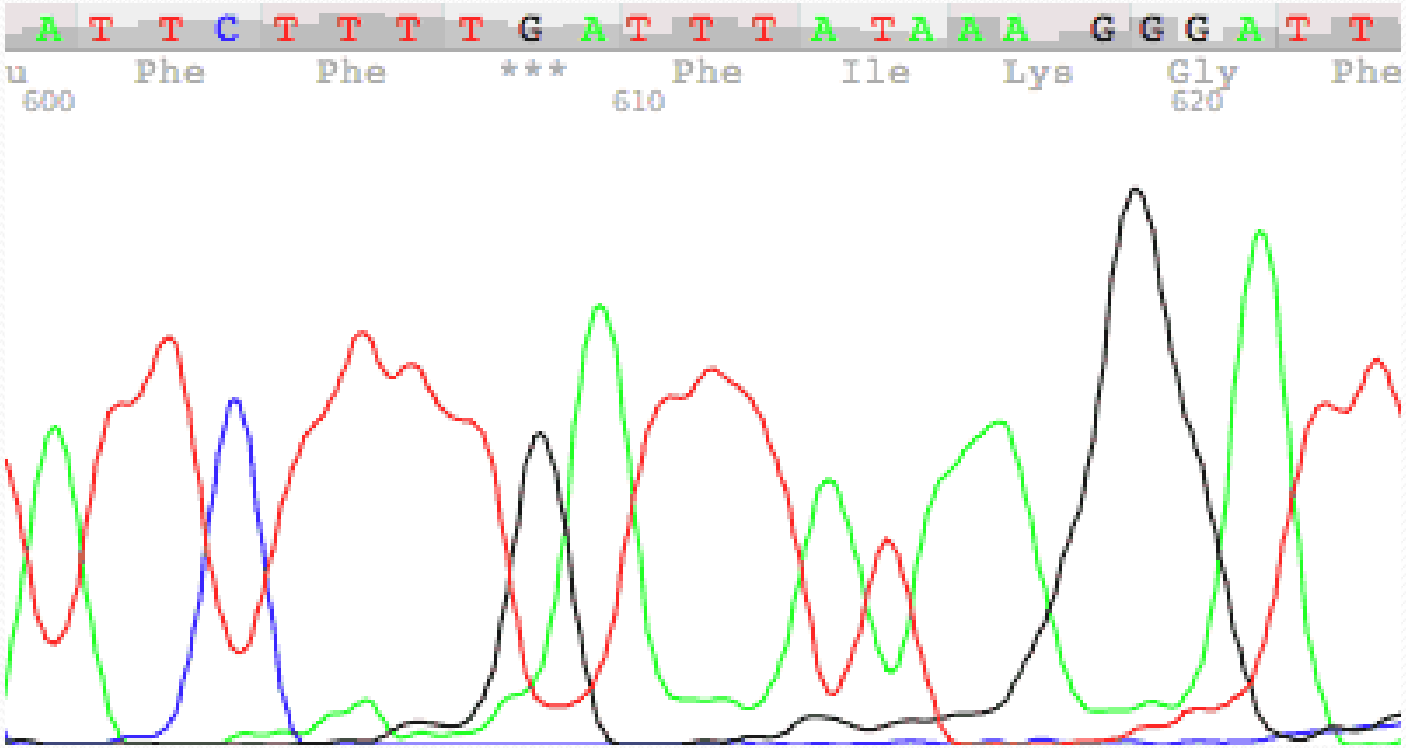
- Two or more templates were present in the reaction
- Two primers were present in the sequencing reactions
- Two priming sites are present in DNA template
- Different sequencing reactions were accidentally mixed at the clean up stage

Short DNA sequencing read lengths



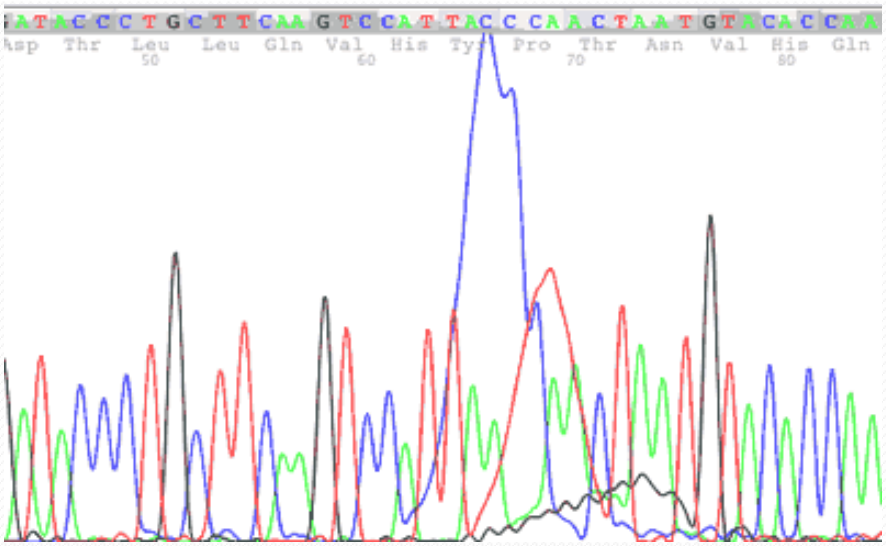
- Too much template DNA
- Excessive dilution of the BigDye reagent
- Too much primer

Blurry trace chromatogram peaks

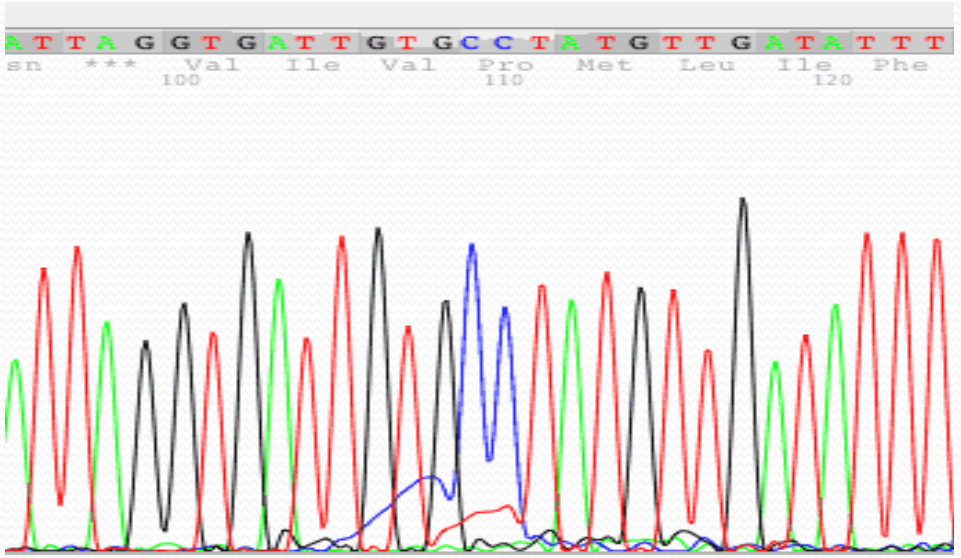


- **Capillary overload** (dirty samples with large amount of DNA, proteins or salts)
- **High sequencing run voltages**

Dye Blobs (excess free dye terminators)



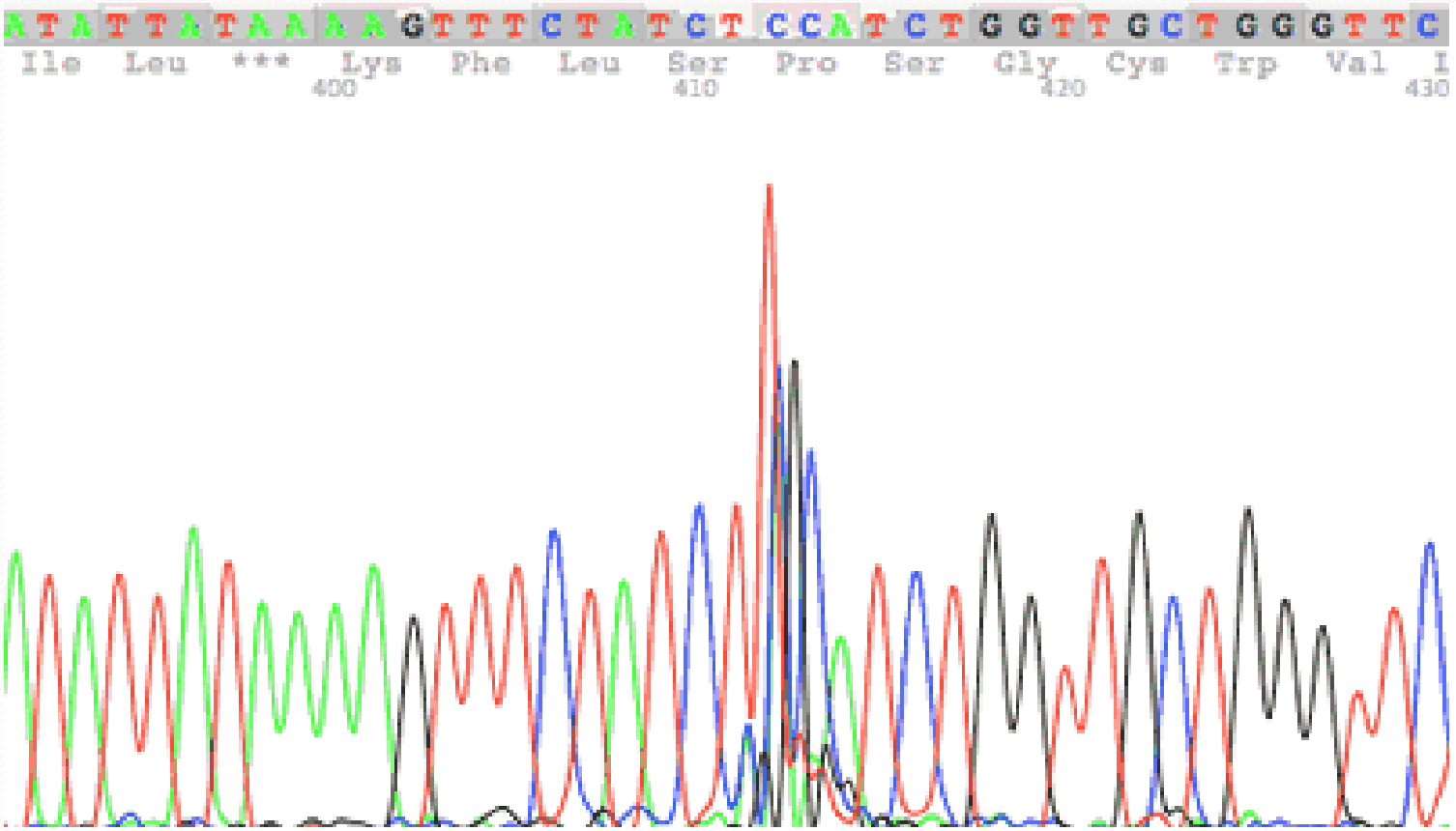
Moderate dye blob



Minor dye blob

-Poor post-sequencing clean up
-Lost pellet
-Failed reaction

Sharp trace signal spikes

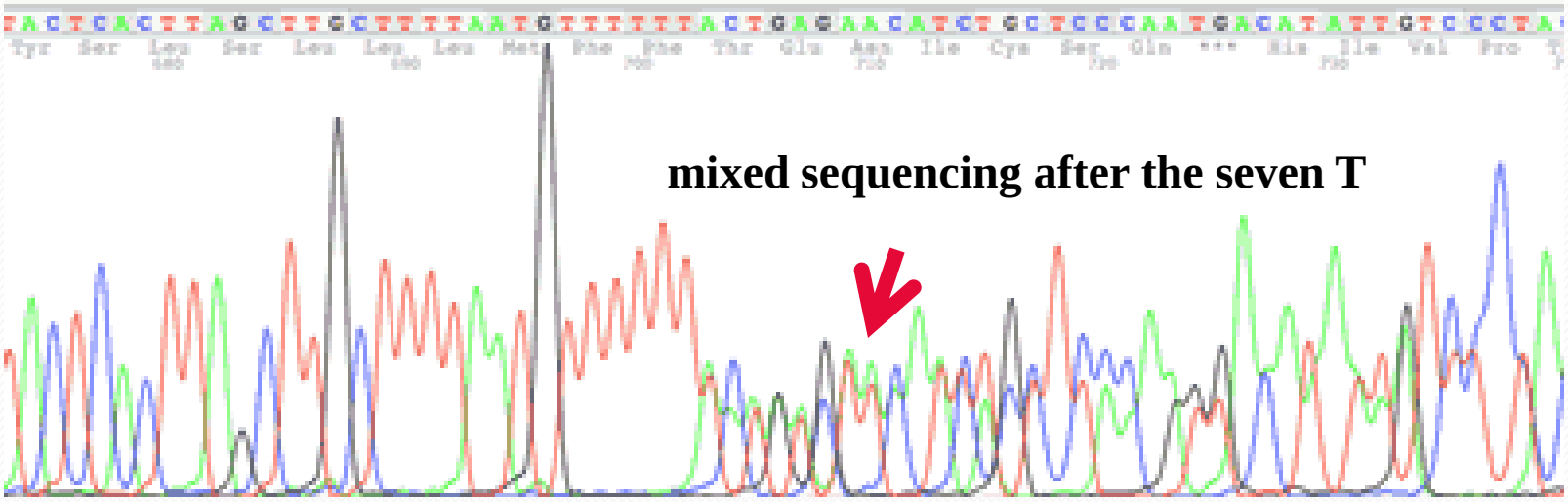


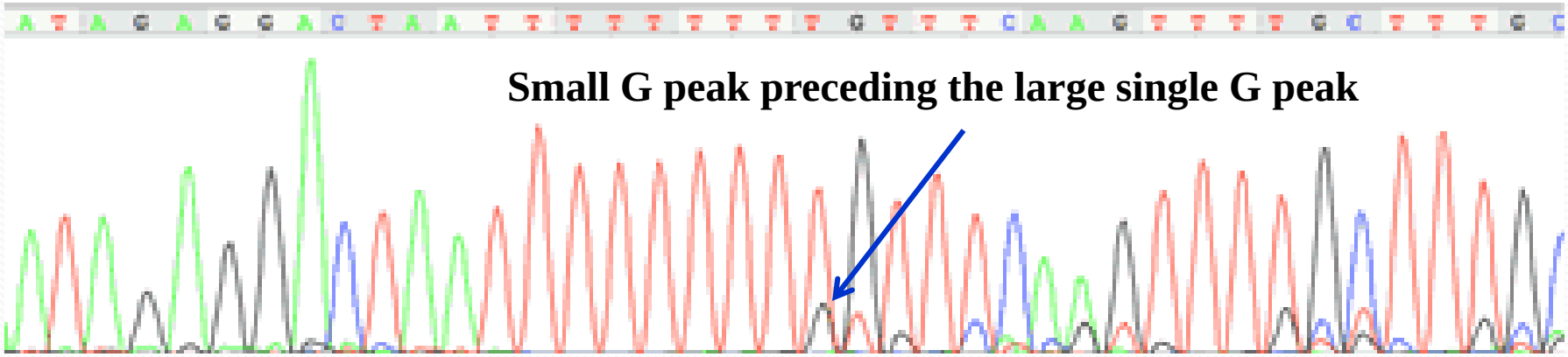
Small gas bubbles forming in the capillaries during electrophoreses

Sequence Slippage

Long runs of a mono nucleotide base

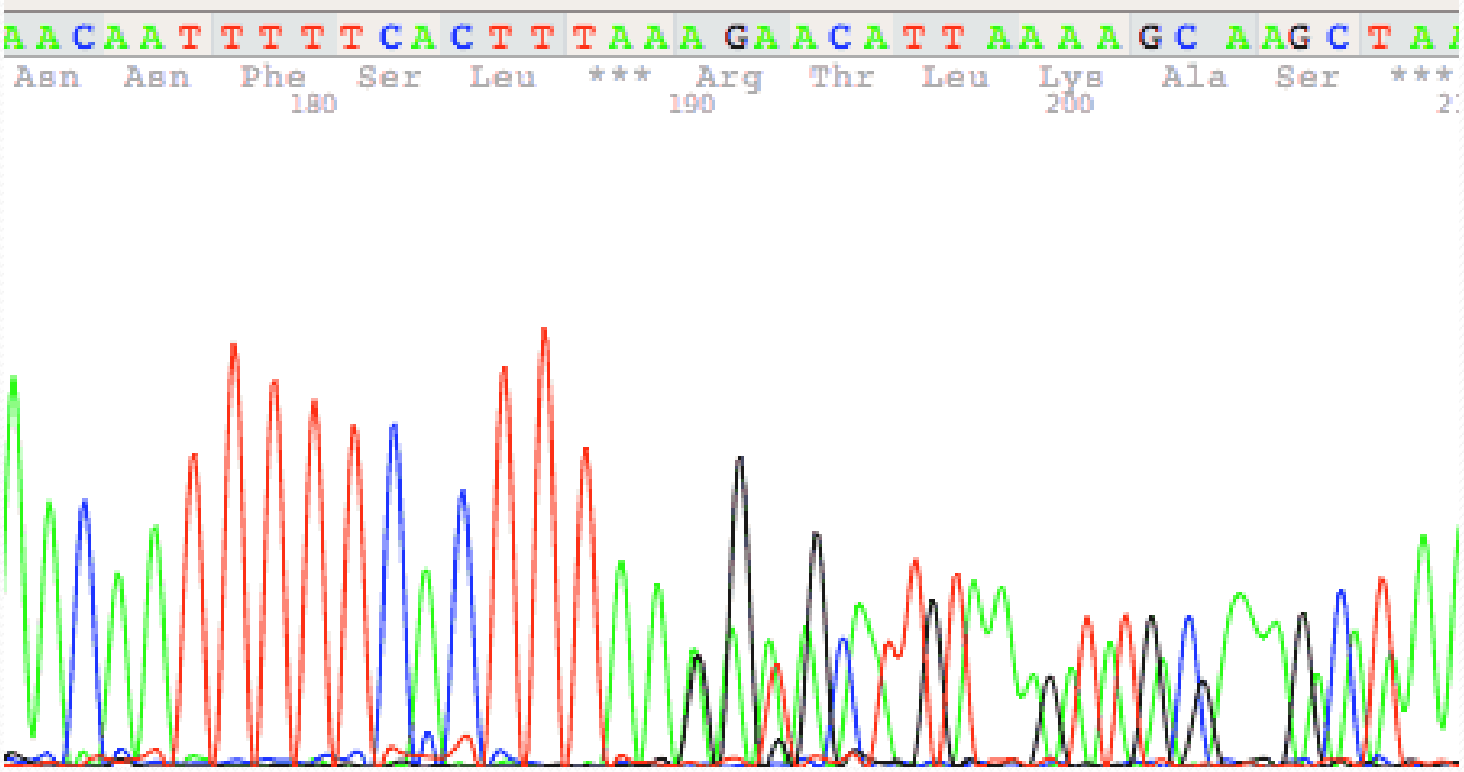
Template insertion and deletions





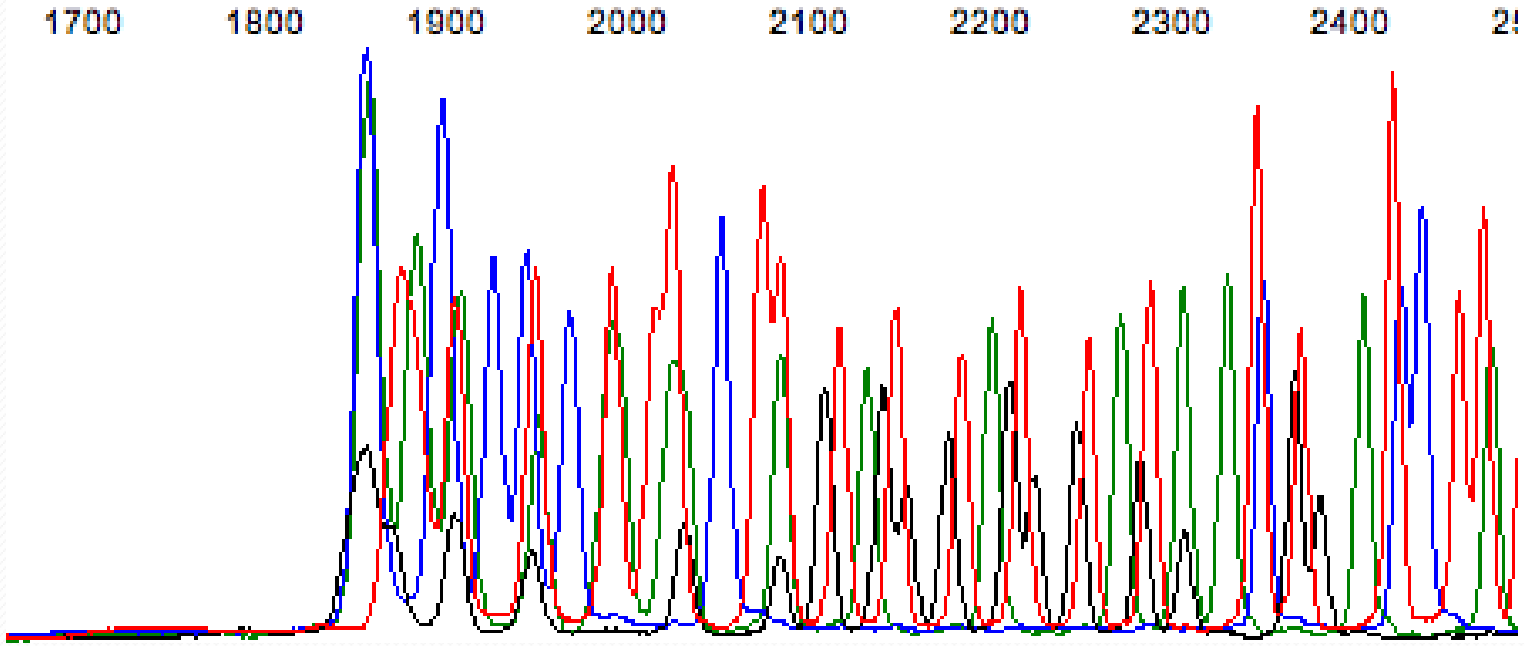
- Direct sequencing of PCR products derived from polymorphic templates
- Random mutation has occurred during the cloning process

Chimeras and sequence rearrangements



- Cloning two or more unrelated DNA fragments at once in the same vector
- 'Unstable' sequences like repeats as long poly A tails

Delayed sequencing signal start



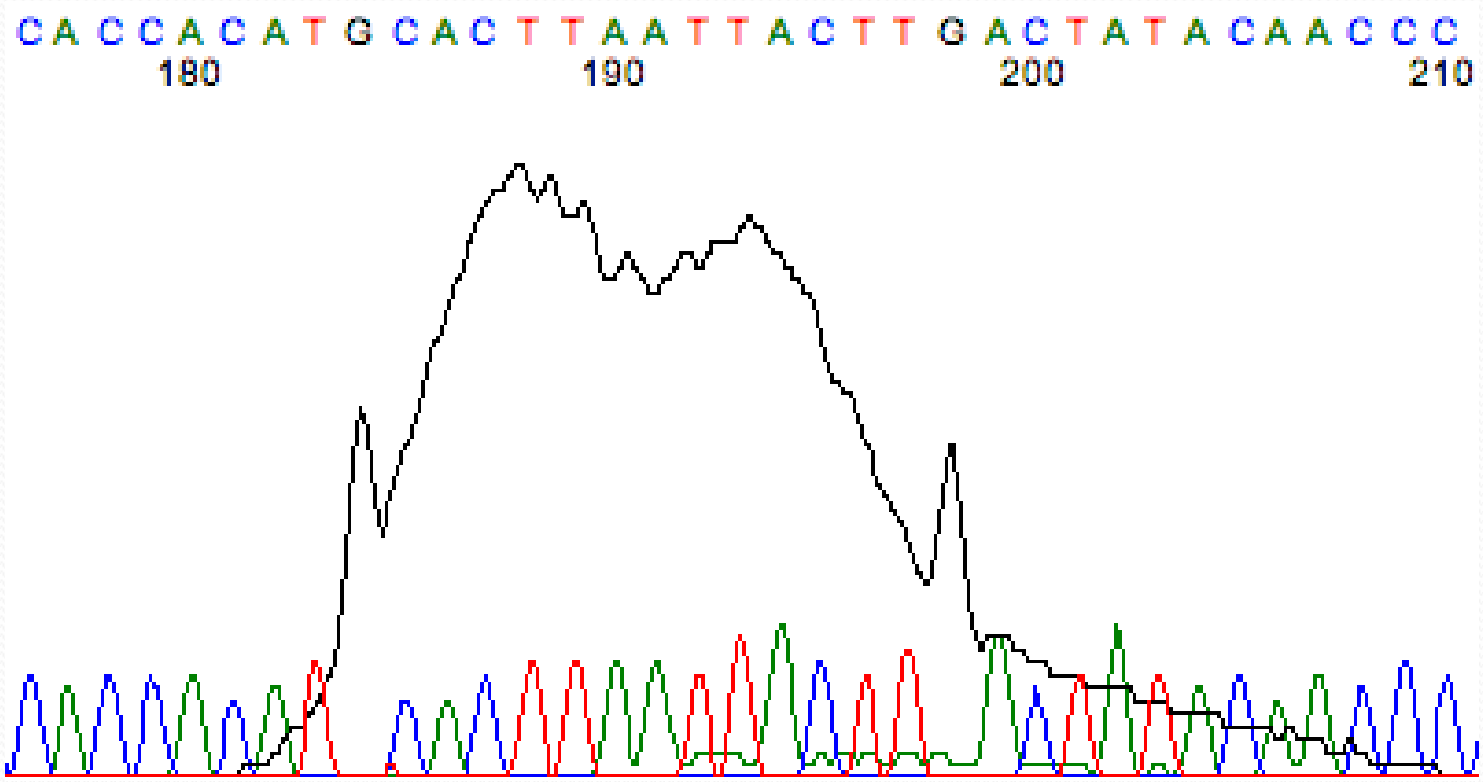
- Capillary overload

The two most common causes of capillary overload are:-

too much template DNA

dirty template DNA contaminated with proteins an/or salt

Late "G" dye peaks



The breakdown of the BigDye sequencing chemistry before sample clean up (excess heating or high numbers of freeze/ thaw cycles)

What can sequence tell us?

-Knowledge about the physical nature of genome

- Classification of microorganisms

- Relationship among diverse genomes

- Information about

  - the origin of microorganism

  -its movement in different population

-Searching for restriction endonuclease cleavage sites

- Vectors designed to over-express proteins or to generate protein

-Identification of protein-coding regions (ORF) within the DNA sequence

Similarity of DNA and protein data bases, can lead to important insights about the function and structure of a cloned gene and its product

-Analysis of the 5′ and 3′ non-coding regulatory regions of a gene.

-DNA sequence information is essential for site directed mutagenesis

-DNA sequence information (sequence tagged sites, or STS; or expressed sequence tags, or ESTs) are the basis of methods for mapping and ordering large DNA segments cloned into yeast or bacterial artificial chromosomes (CYACs; BACs) or cosmids


THANK YOU