SNP Genotyping Detection Methods

Prof. Dr. Batool Hassan AL-Ghurabi

Biological Background

- How can researchers hope to identify and study all the changes that occur in so many different diseases?
- How can they explain why some people respond to treatment and not other?
- Answer of these question can be given by SNP

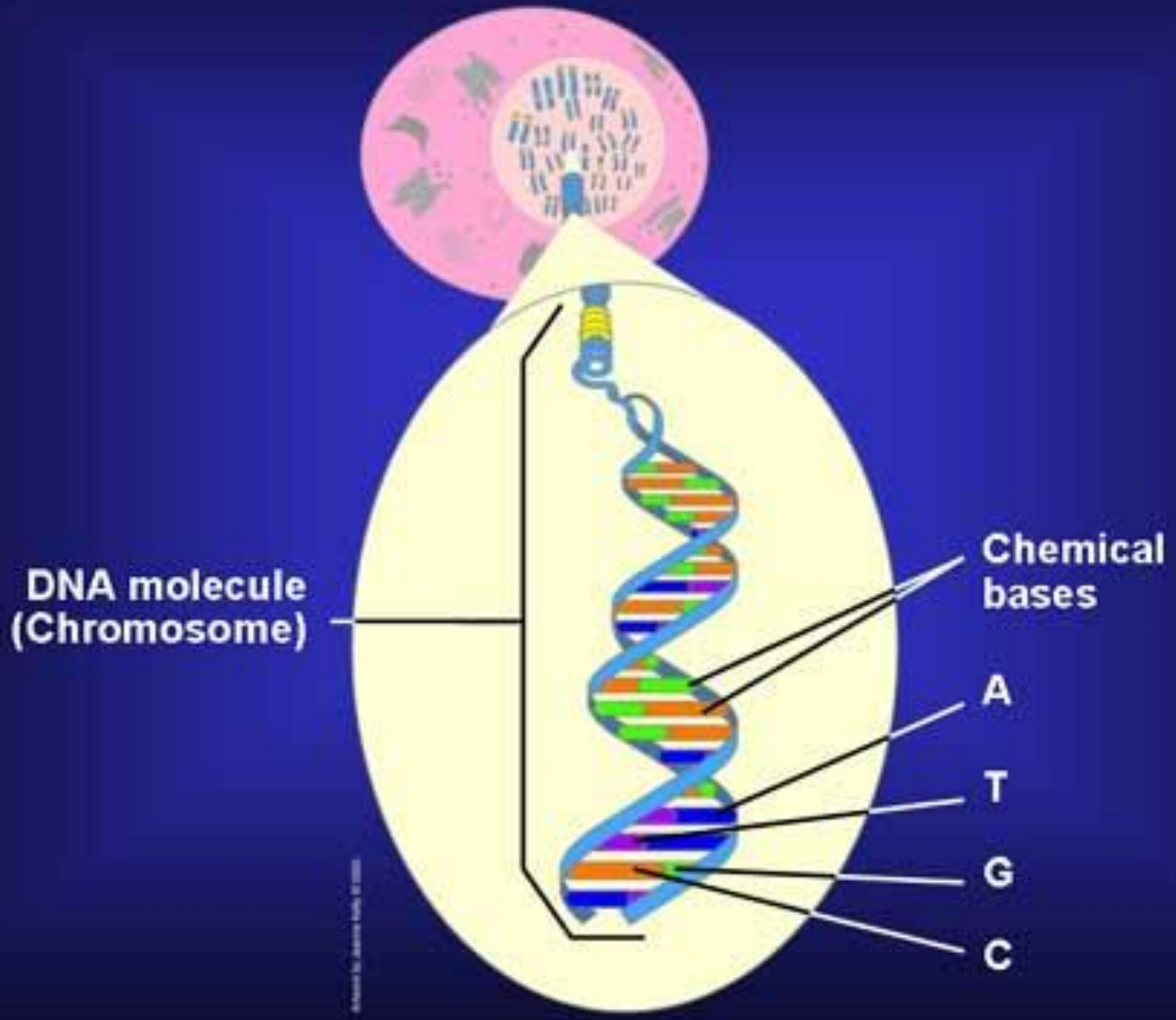
Single nucleotide polymorphism

- A Single Nucleotide Polymorphism, also known as Simple Nucleotide Polymorphism, is a DNA sequence variation occurring commonly within a population (e.g. 1%) in which a single nucleotide — A, T, C or G — in the genome differs between members of a biological specie .
- Pronounced snips

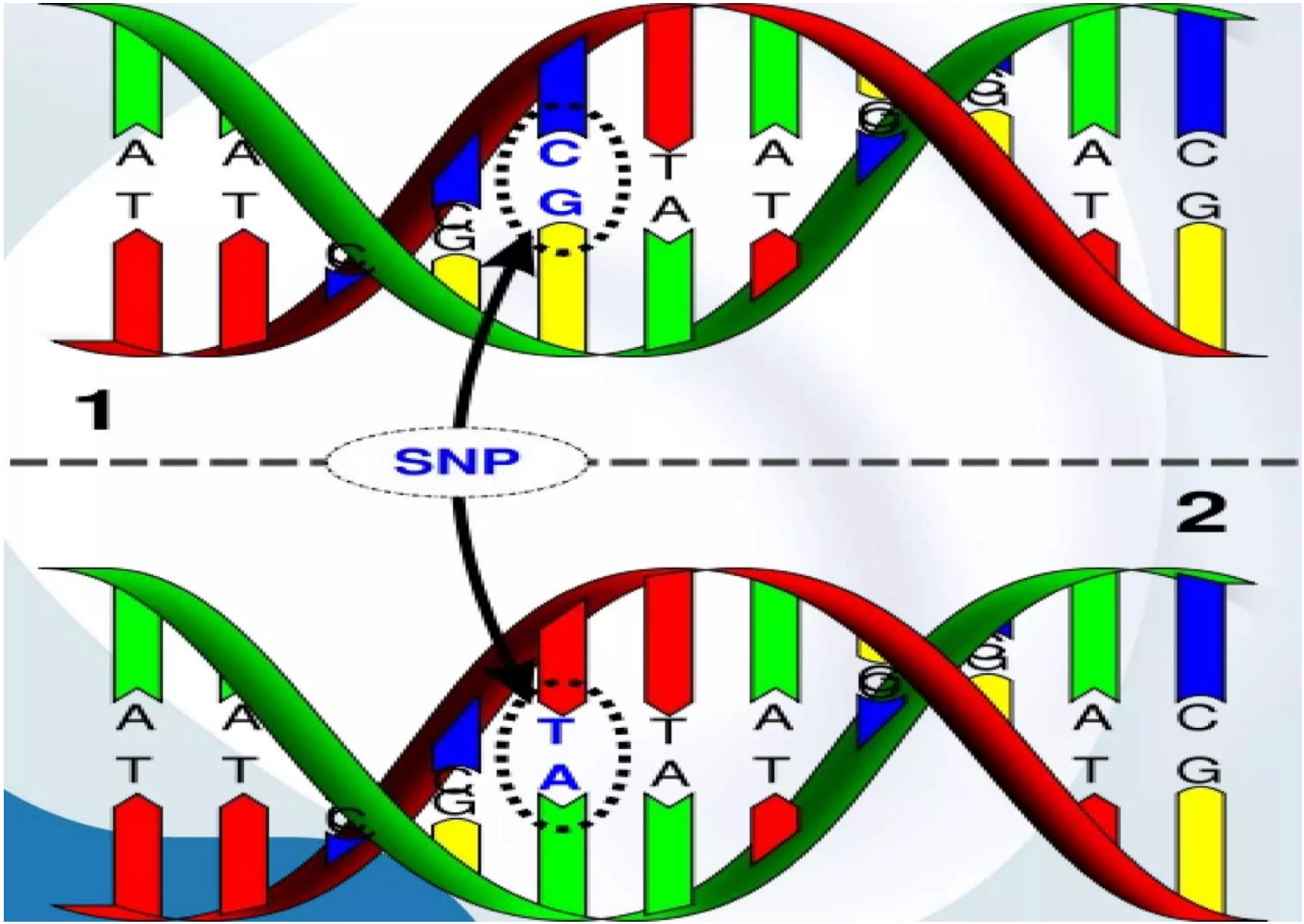
Common type of genetic variation among people

- Each SNP represents a difference in a single DNA building block called as nucleotide

DNA and Chromosome Structure



NATIONAL
CANCER
INSTITUTE


1
SNP
2

Difference between SNP and Mutation

SNP	Mutation
SNP is a change in the single-nucleotide of a genome. Also, it is a type of mutation.	**Mutation is the variation in DNA base pairs caused due to insertion, deletion, duplication or substitution of base pairs.**
The variation is seen only in a single nucleotide.	**The variation can be due to changes in many or even a single nucleotide.**
The SNP variation is available in a minimum of 1% of the population.	**The mutation frequency is available in less than 1% of the population.**
Example – In the sequence ATAGC, the substitution of G by C will produce ATACC. This change in a single nucleotide is termed as SNP.	**The mutations are of different types. The missense mutation, silent mutation and nonsense mutation are some of them.**

Some Facts

- In human beings, 99.9 percent bases are same.
- Remaining 0.1 percent makes a person unique.
  - Different attributes / characteristics / traits
    - how a person looks,
    - diseases he or she develops.

- These variations can be:
  - Harmless (change in phenotype)
  - Harmful (diabetes, cancer, heart disease, Huntington's disease, and hemophilia)
  - Latent (variations found in coding and regulatory regions, are not harmful on their own, and the change in each gene only becomes apparent under certain conditions e.g. susceptibility to lung cancer)

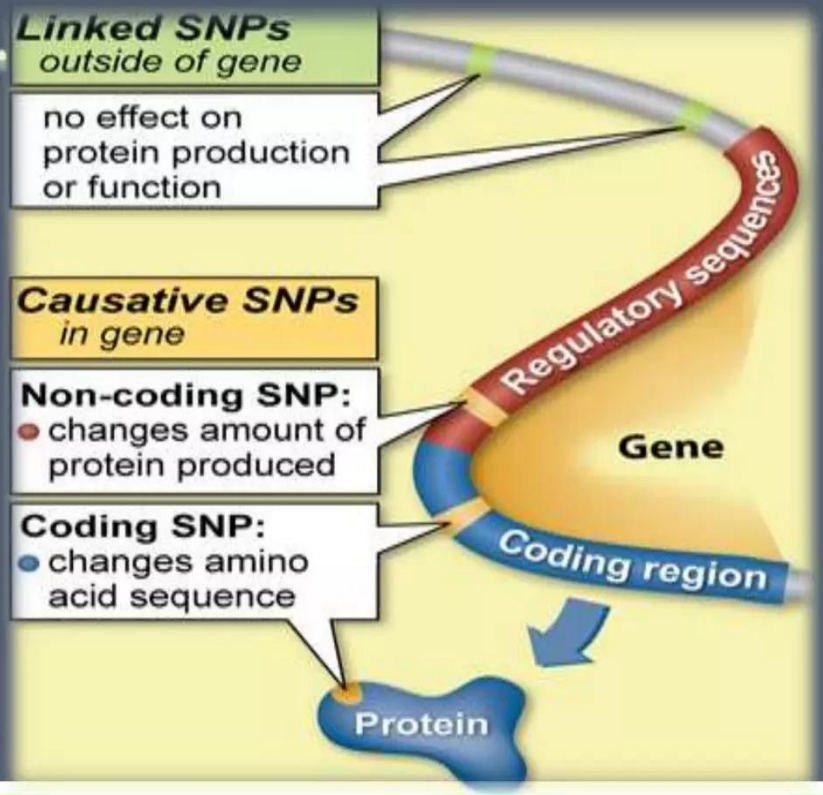
SNPs facts

- SNPs are found in
  - coding and (mostly) non coding regions.
- Occur with a very high frequency
  - about 1 in 1000 bases to 1 in 100 to 300 bases.
- The abundance of SNPs and the ease with which they can be measured make these genetic variations significant.
- SNPs close to particular gene acts as a marker for that gene.
- SNPs in coding regions may alter the protein structure made by that coding region.

Types of SNPs

- **Linked SNPs** :Or called indicative SNPs They do not reside within gene and do not effect protein function. Nevertheless they do correspond to a particular drug response
- **Causative SNPs** :Affect the way of a protein function,corelating with the disease or influencing a person's response to medication
- **Non-coding region** :A segment of DNA that does comprise a gene and thus does not code for a protein.
- **Coding region:** regions of DNA or RNA sequence that codes for protein

- Coding region divided into
- Synonymous SNP: The substitution of one base for another in an exon of a gene coding for a protein such that the produced amino acid sequence is not modified.
- Non synonymous SNp:A NON -SYNONYMOUS SUBSTITUTION RESULTS IN A CHANGE IN AMINO ACID.



SNP MAPPING

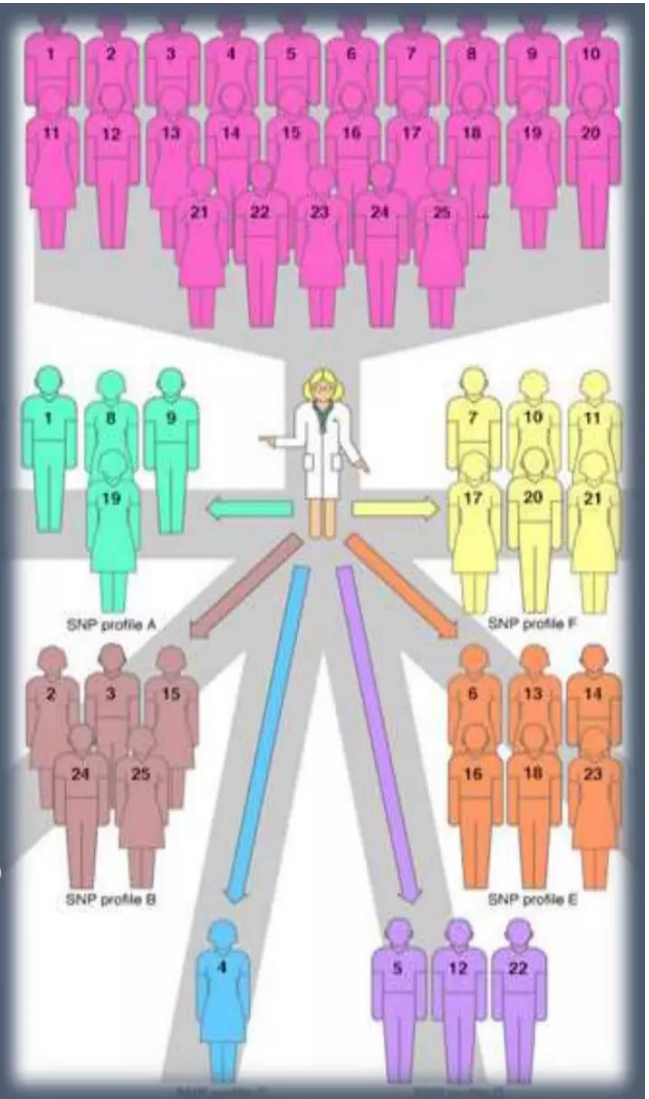
- Sequence genomes of a large number of people
- Compare the base sequences to discover SNPs.
- Generate a single map of the human genome containing all possible SNPs

International Hap Map

- Empirical genotype data from > 3 million SNPs in a limited samples of 270 individuals from 4 populations
  - European origin
  - West African
  - Chinese
  - Japanese

SNP Profiles

- Genome of each individual contains distinct SNP pattern.
- People can be grouped based on the SNP profile.
- SNPs Profiles important for identifying response to Drug Therapy.
- Correlations might emerge between certain SNP profiles and specific responses to treatment.



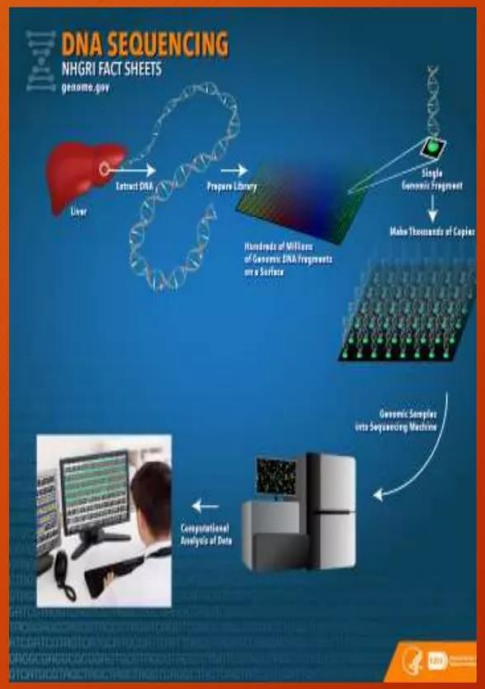
SNP analysis

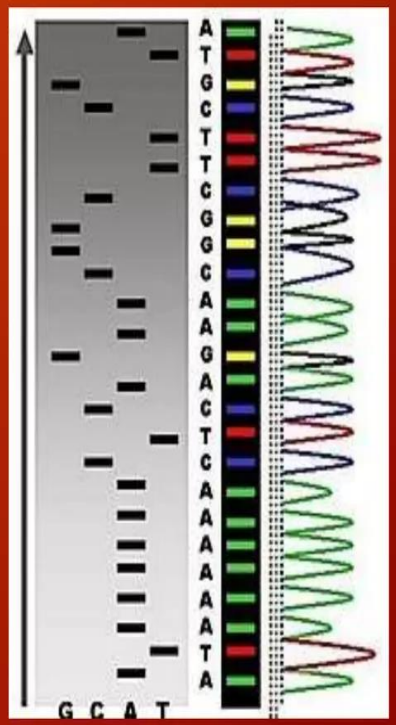
Analytical methods to discover novel SNPs and detect known SNPs include:

- DNA sequencing
- mass spectrometry
- single-strand conformation polymorphism (SSCP)
- electrochemical analysis
- capillary electrophoresis
- denaturating HPLC and gel electrophoresis
- restriction fragment length polymorphism

Technology for Targeted SNP Discovery

- Targeted SNP discovery focuses on most and SNP in regions of interest.
- Recently many SNP scanning technologies are developed.
- Sequencing is the perfect method but not cost effective.





Enzyme-based methods

- A broad range of enzymes including DNA ligase, DNA polymerase and nucleases have been employed to generate high-fidelity SNP genotyping methods.

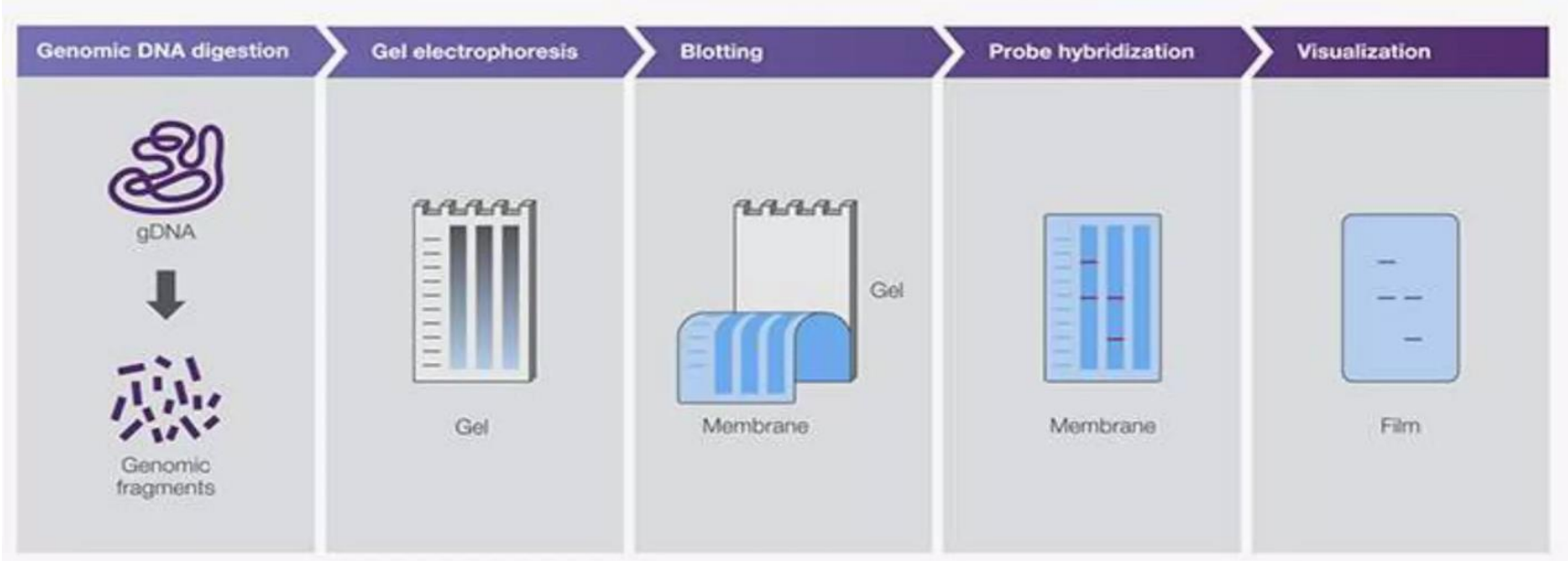
Restriction fragment length polymorphism

- Restriction fragment length polymorphism (RFLP) is the simplest and earliest method to detect SNPs.
- SNP-RFLP use different restriction endonucleases enzymes with high affinity to unique and specific restriction sites.
- By a digestion on a genomic sample and determining fragment lengths through a gel assay it is possible to ascertain whether or not the enzymes cut the expected restriction sites.

Restriction fragment length polymorphism

- A failure to cut the genomic sample results in an larger than expected fragment implying that there is a mutation at the point of the restriction site which is rendering it protection from nuclease activity.
- RFLP a poor choice for high throughput analysis.

RFLP- *"rif-lip"*

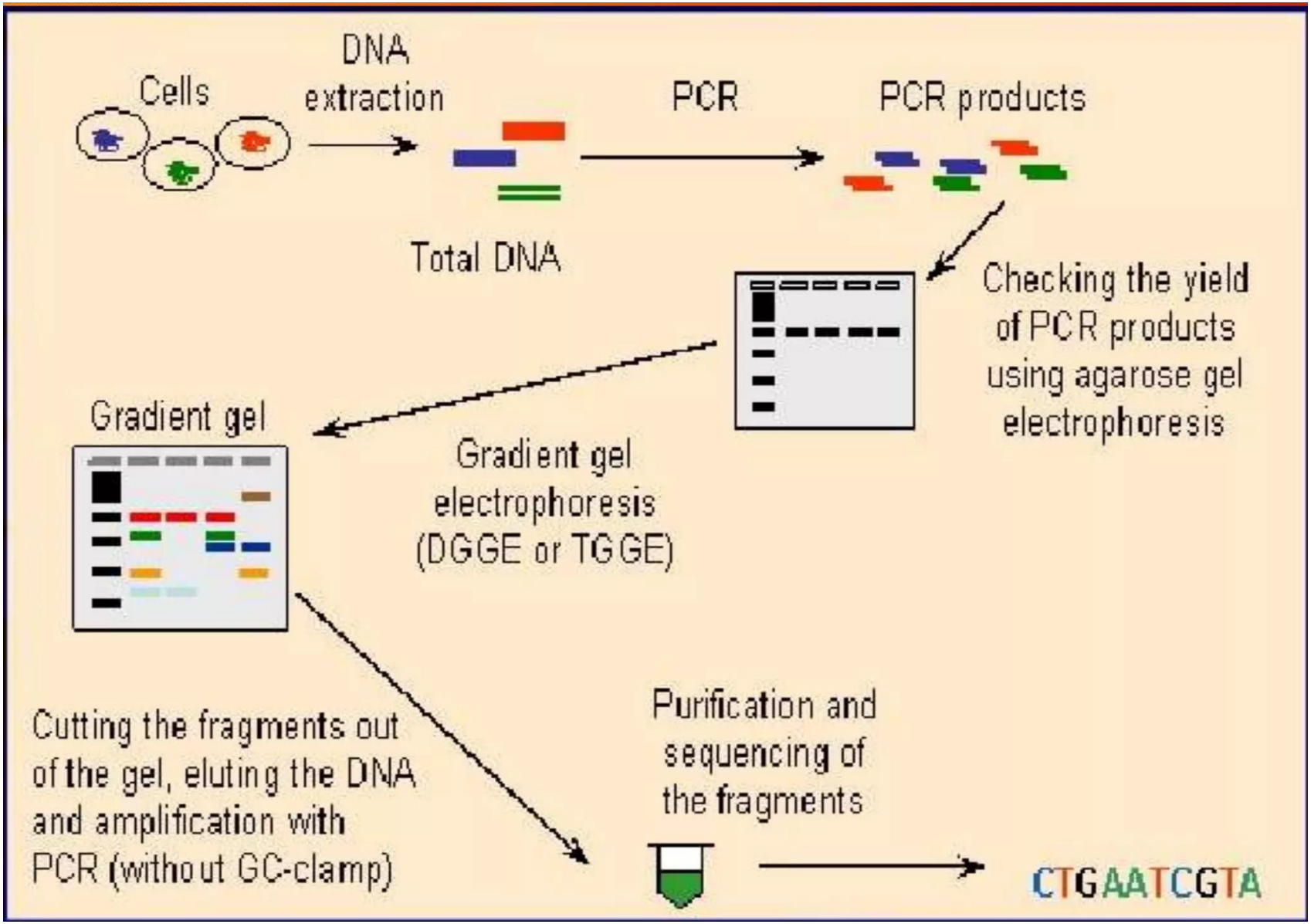


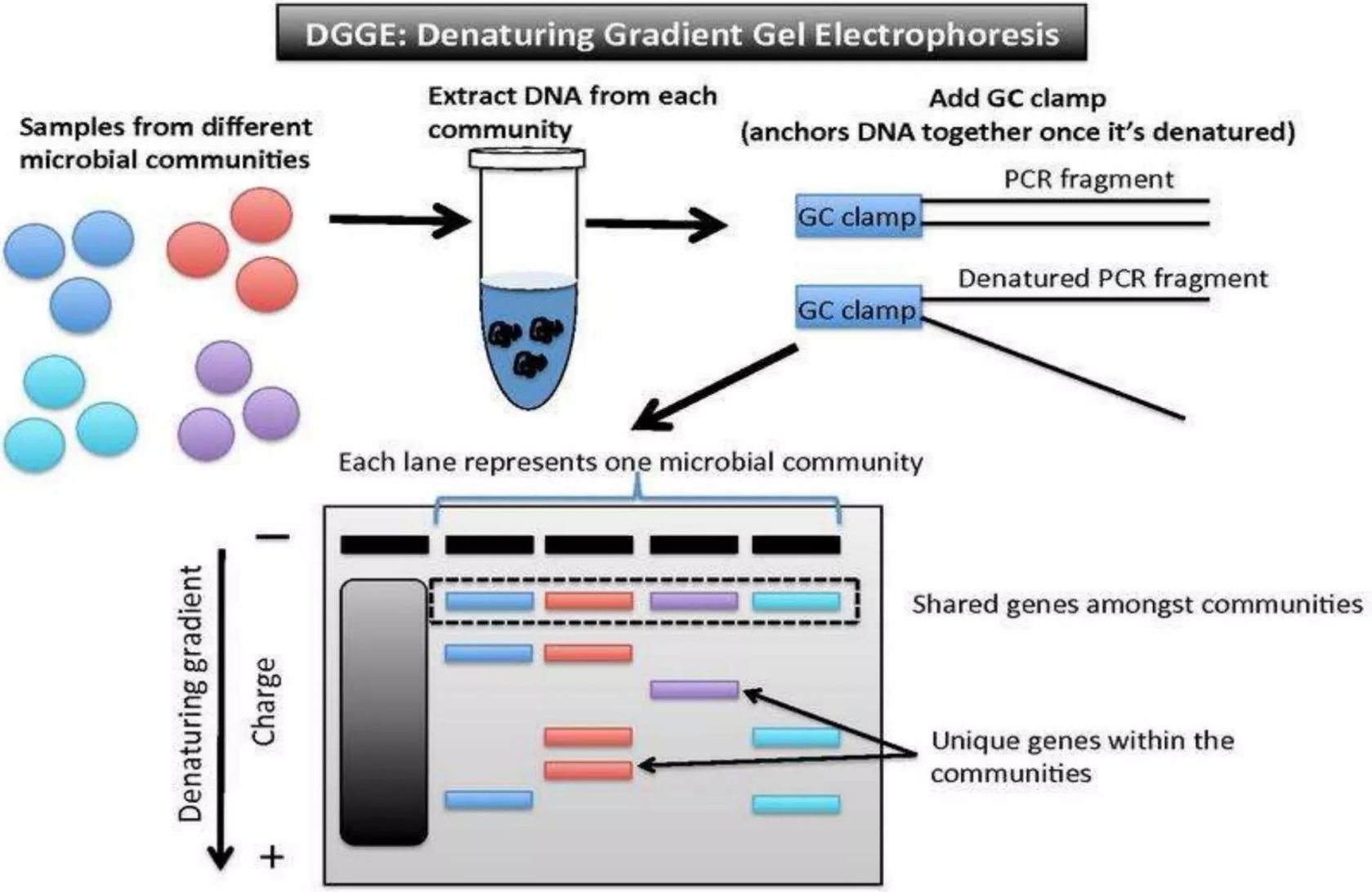
Restriction Fragment Length

Polymorphism

Denaturing Gradient Gel Electrophoresis

- Denaturation is highly depend on DNA sequence.
- In many cases a single nucleotide change the melting point.
- And change melting point shows unique position on gel


Cells
DNA extraction
Total DNA
PCR
PCR products
Checking the yield of PCR products using agarose gel electrophoresis
Gradient gel
Gradient gel electrophoresis (DGGE or TGGE)
Cutting the fragments out of the gel, eluting the DNA and amplification with PCR (without GC-clamp)
Purification and sequencing of the fragments
CTGAATCGTA


DGGE: Denaturing Gradient Gel Electrophoresis
Samples from different microbial communities
Extract DNA from each community
Add GC clamp
(anchors DNA together once it's denatured)
PCR fragment
GC clamp
Denatured PCR fragment
GC clamp
Each lane represents one microbial community
Denaturing gradient
Charge
−
+
Shared genes amongst communities
Unique genes within the communities

Gel electrophoresis

- is a method for separation and analysis of macromolecules and their fragments Nucleic acid molecules are separated by applying an electric field to move the negatively charged molecules through a matrix of agarose or other substances.
- Shorter molecules move faster and migrate farther than longer ones.
- This phenomenon is called sieving.

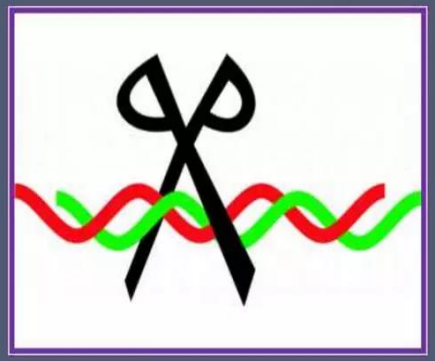
Procedure:

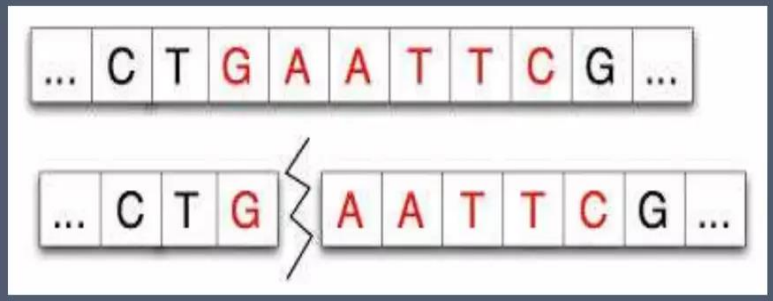
Step 1..place DNA into tubes..

Step 2.. The polymerase chain reaction uses a machine called thermocycler to quickly copy a piece of DNA.

Step 3.. Place restriction

restriction enzyme will cut





- restriction enzyme will cut DNA into different sizes according to its sequnce.

Step 4.. Dye DNA and add into gel

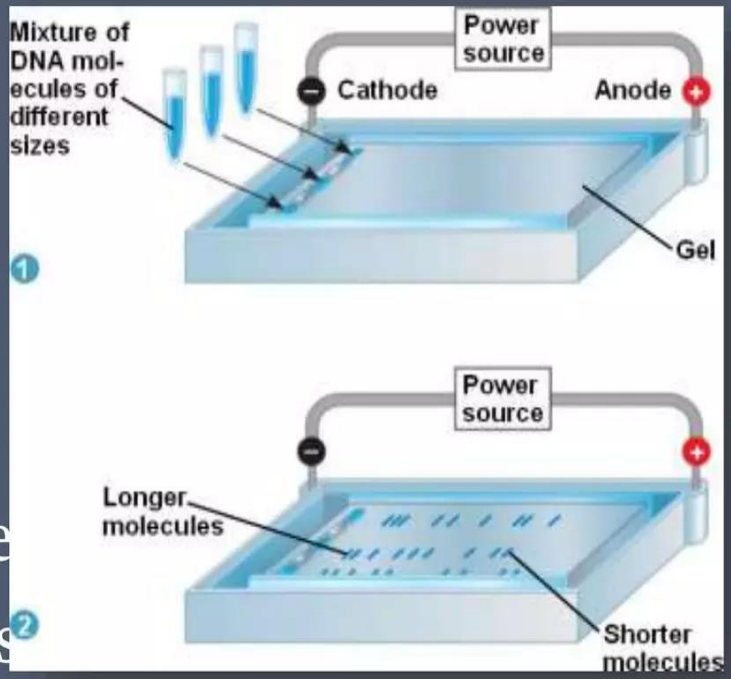
..the gel is made of agarose.

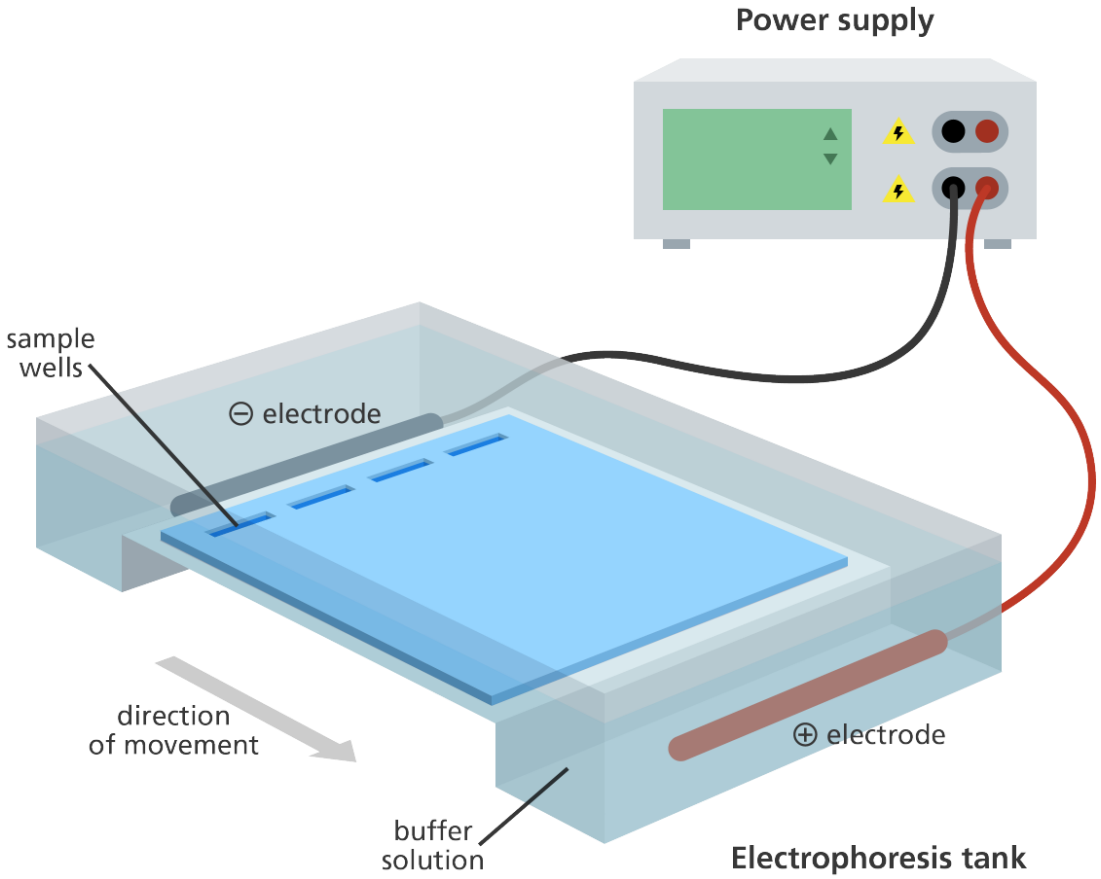
Step 5..

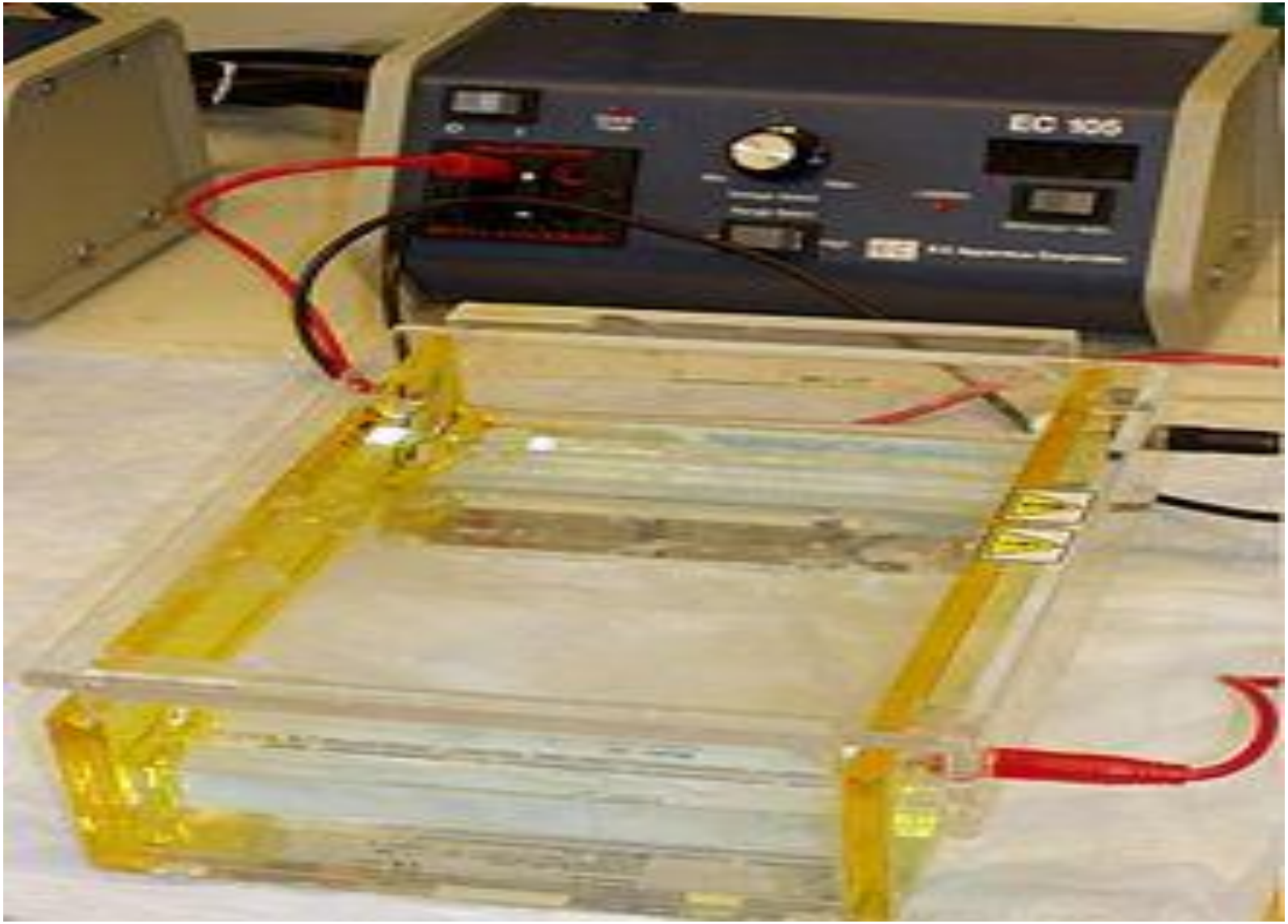
Run electric current through gel.

Step 6..

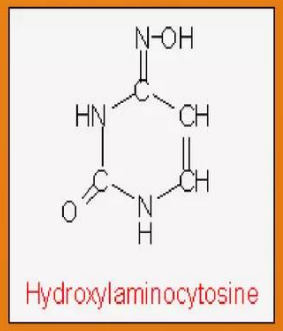
DNA is negatively charged and hence will move towards the positive charged end of the gel.Smaller pieces of DNA travel faster than larger pieces of DNA.




Power supply
sample wells
⊖ electrode
direction of movement
⊕ electrode
buffer solution
Electrophoresis tank


EC 105

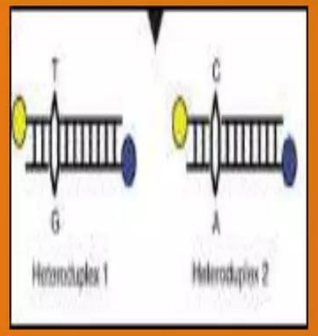
Chemical Cleavage Of Mismatch



Find SNP by specific chemicals

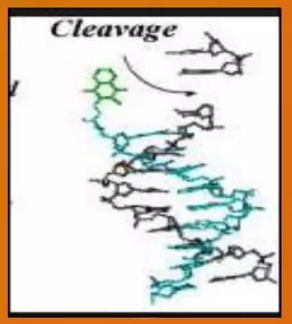
Potassium Permanganate

Hydroxylamine

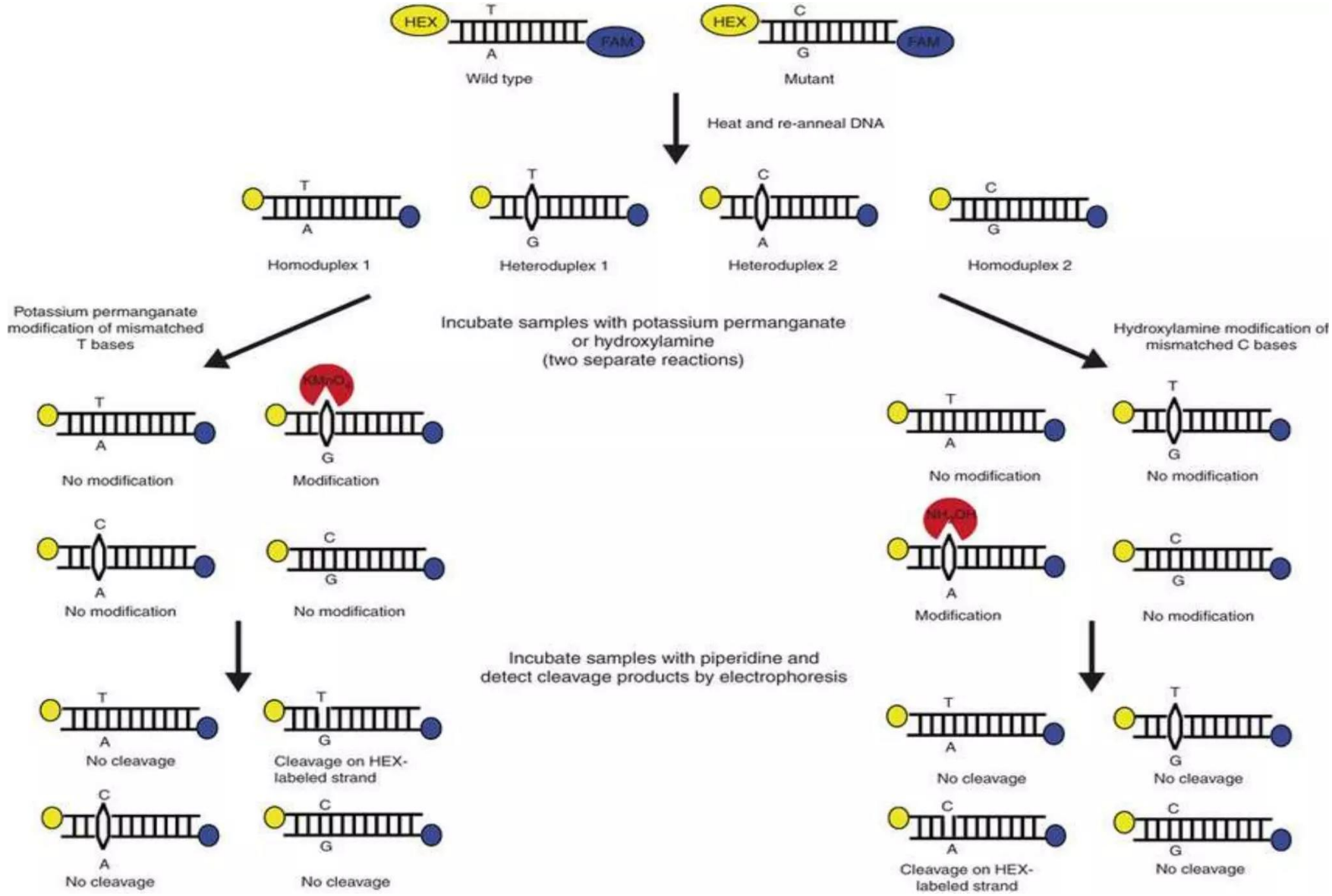


Target sequence is PCR to create heteroduplex

- Denature
- Reanneal



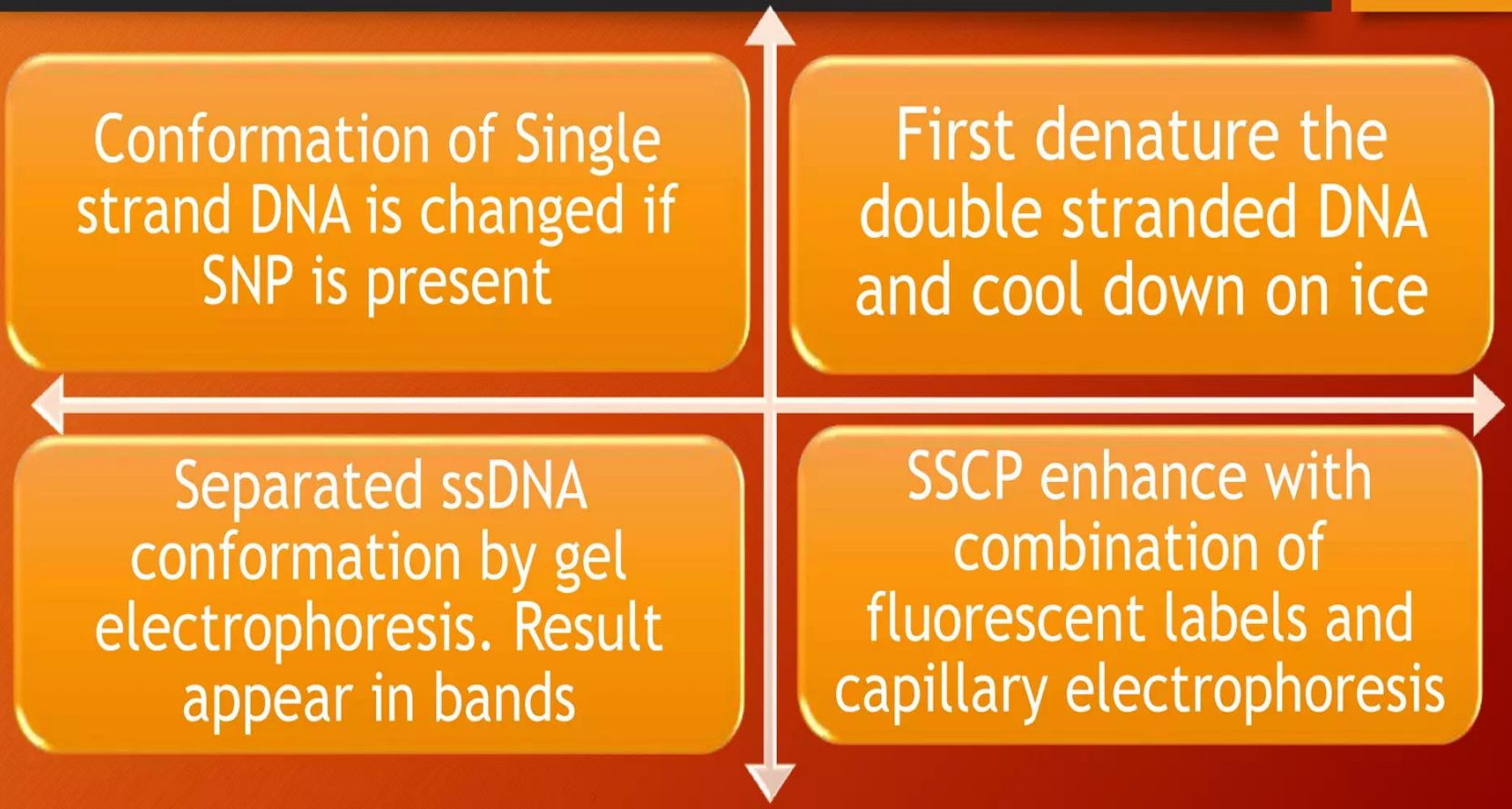
Treated with respective chemicals and cleaved with piperidine and run on gel


HEX
T
A
FAM
Wild type
HEX
C
G
FAM
Mutant
Heat and re-anneal DNA
T
A
Homoduplex 1
T
G
Heteroduplex 1
C
A
Heteroduplex 2
C
G
Homoduplex 2
Potassium permanganate modification of mismatched T bases
Incubate samples with potassium permanganate or hydroxylamine (two separate reactions)
Hydroxylamine modification of mismatched C bases
KMnO4
No modification
Modification
No modification
No modification
NH2OH
Modification
No modification
No modification
Incubate samples with piperidine and detect cleavage products by electrophoresis
No cleavage
Cleavage on HEX-labeled strand
No cleavage
No cleavage
No cleavage
No cleavage
Cleavage on HEX-labeled strand
No cleavage

Mass spectrometry

- may be used to determine DNA sequences.
- mass spectrometry has specifically been investigated as an alternative method to gel electrophoresis for visualizing DNA fragments. With this method, DNA fragments are compared by mass rather than by size. The mass of each nucleotide is different from the others and this difference is detectable by mass spectrometry.
- Single-nucleotide mutations in a fragment can be more easily detected with MS than by gel electrophoresis alone

Single-stranded Conformation Polymorphism (SSCP)

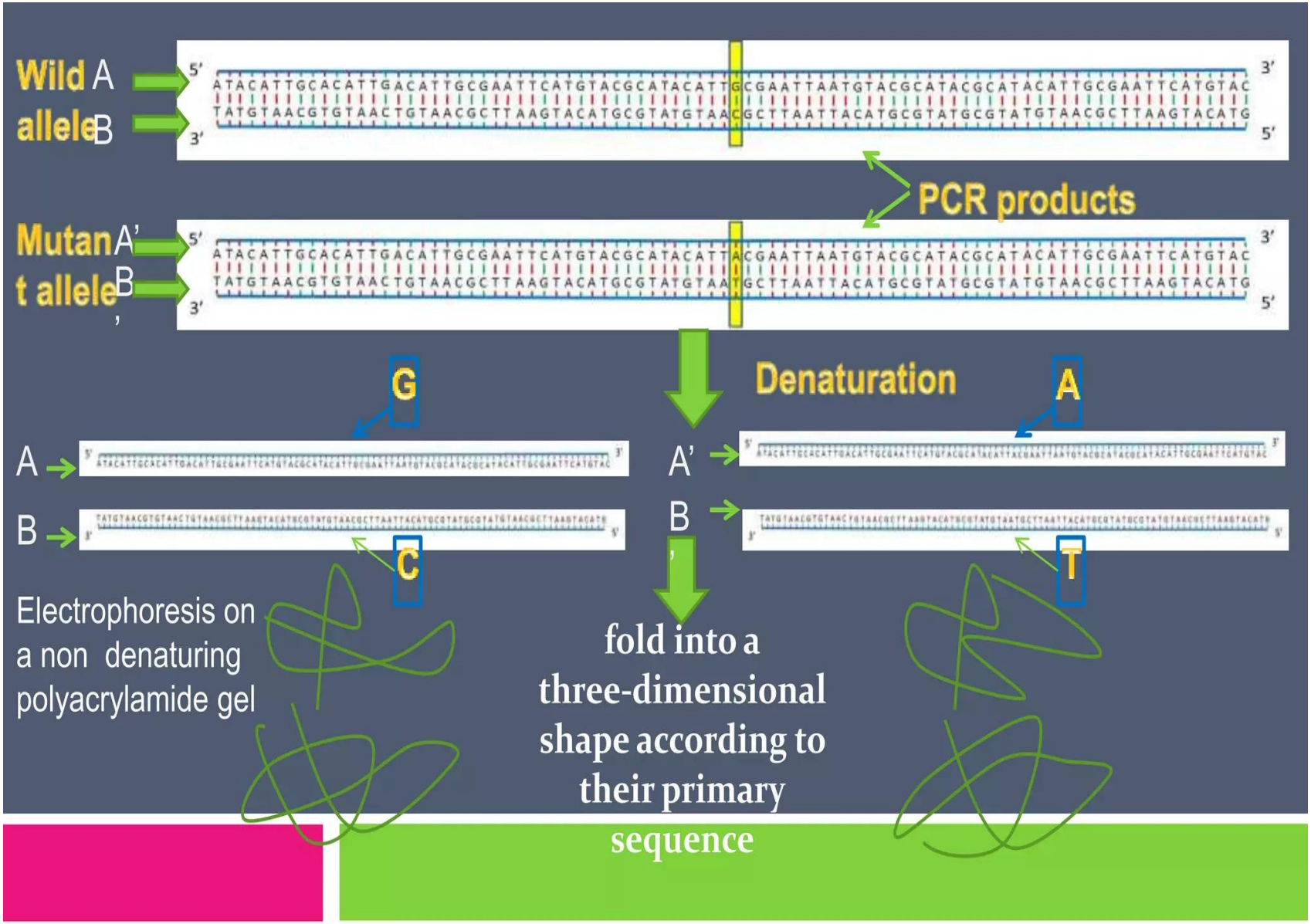


Instruments of SSCP

DNA Isolation kit

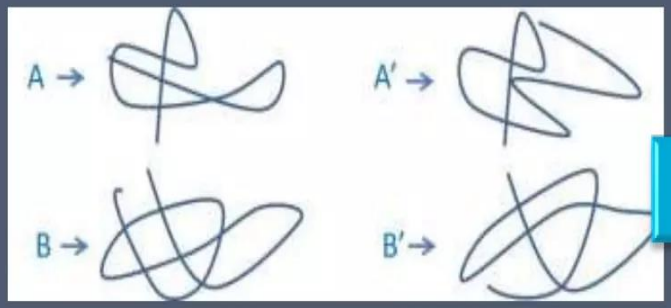
Amplify region of interest by PCR using fluorescent labeled primer.

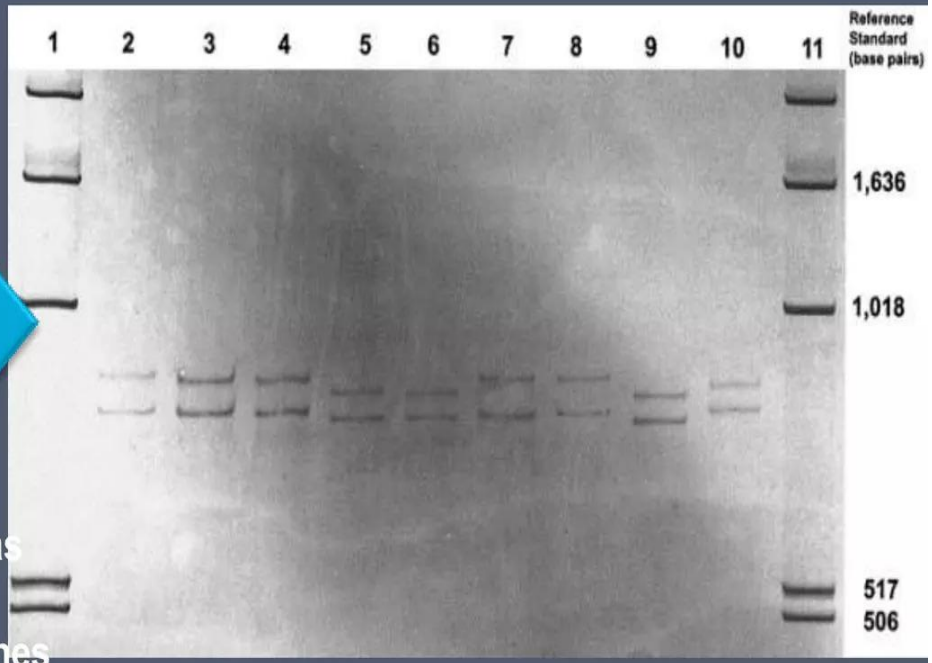
Capillary electrophoresis


Wild allele
A
B
Mutant allele
A'
B'
5'
3'
ATACATTGCACATTGACATTGCGAATTCATGTACGCATACATTGCGAATTAATGTACGCATACGCATACATTGCGAATTCATGTAC
TATGTAACGTGTAACTGTAACGCTTAAGTACATGCGTATGTAACGCTTAATTACATGCGTATGCGTATGTAACGCTTAAGTACATG
PCR products
Denaturation
G
A
C
T
A
A'
B
B'
Electrophoresis on a non denaturing polyacrylamide gel
fold into a three-dimensional shape according to their primary sequence

Wild type

Mutant type





Sample SSCP Gel Result and Interpretation. **DNA was isolated and amplified SCCP analysis of the DNA shows sets of alleles usually inherited as a unit. Lanes 3 and 4 were identical haplotypes from two individuals. The difference in band migration in adjacent lanes is associated with the number of nucleotide differences (in parentheses): lanes 2-3 (2), lanes 3-4 (0), lanes 4-5 (3), lanes 5-6 (1), lanes 6-7 (3), lanes 7-8 (1), lanes 8-9 (1), and lanes 9-10 (4).**

MutS Protein-binding Assays

MutS Protein

- Are the key element in mismatch repair
- It attach with DNA and activate MutS-H endonuclease that cut & repair DNA

MutS protein-Binding Assay

- MutS immobile in magnetic beads to capture hetroduplex DNA
- DNA is labeled by Biotin and detected by ELISA

Steps of MutS protein-Binding Assay

Step 1	Step 2	Step 3	Step 4
Prepared DNA region with SNP using PCR Amplification	Prepare DNA heteroduplex	Formation of MutS-DNA complex	Detection of MutS-DNA Complex

DNA mismatch-binding proteins

- DNA mismatch-binding proteins can distinguish single nucleotide mismatches and thus facilitate differential analysis of SNPs.

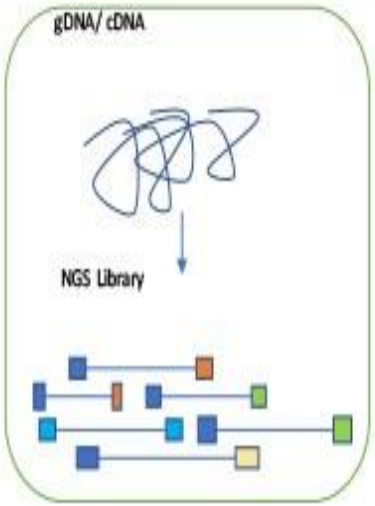
Next-generation sequencing

- Next-generation sequencing (NGS) is a technology for determining **the sequence of DNA or RNA** to study **genetic variation associated** with diseases or other **biological** phenomena.
- Introduced for commercial use in 2005, this method was initially called “massively-parallel sequencing”, because it enabled the **sequencing** of **many DNA strands** at the same time, instead of one at a time as with traditional Sanger sequencing by capillary electrophoresis (CE).

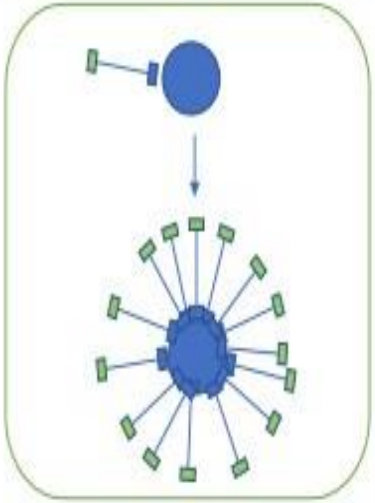
steps of next generation sequencing (NGS)?

- Step 1- Nucleic Acid Extraction and Isolation.
- Step 2- Library Preparation. ...
- Step 3- Clonal Amplification and Sequencing. .
- Step 4 -Data Analysis Using Bioinformatics.

1. Construct Library



2. Clonal Amplification



3. Sequence Library

4. Analyze Data

Advantages of snps

- SNPs are the most frequent form of variation
- They are disease causing mutations in many genes
- They are abundant and slow mutation rates
- May work as the nxt generation of genetic markers

MEDICATION & DIAGNOSIS AT INDIVIDUAL LEVEL

- IN DISEASE DIAGNOSIS
- IN DRUG DISCOVERY & DEVELOPMENT
- IN DRUG RESPONSE
- INVESTIGATION OF MIGRATION PATTERNS

Disadvantages of snp

Less alleles

–Each marker is less informative

Therefore have to genotype many more SNPs to get same level of information about DNA sample

A wide range of disease occurred by this in humans

- Sickle cell anemia
- B thalasemia
- diabetes

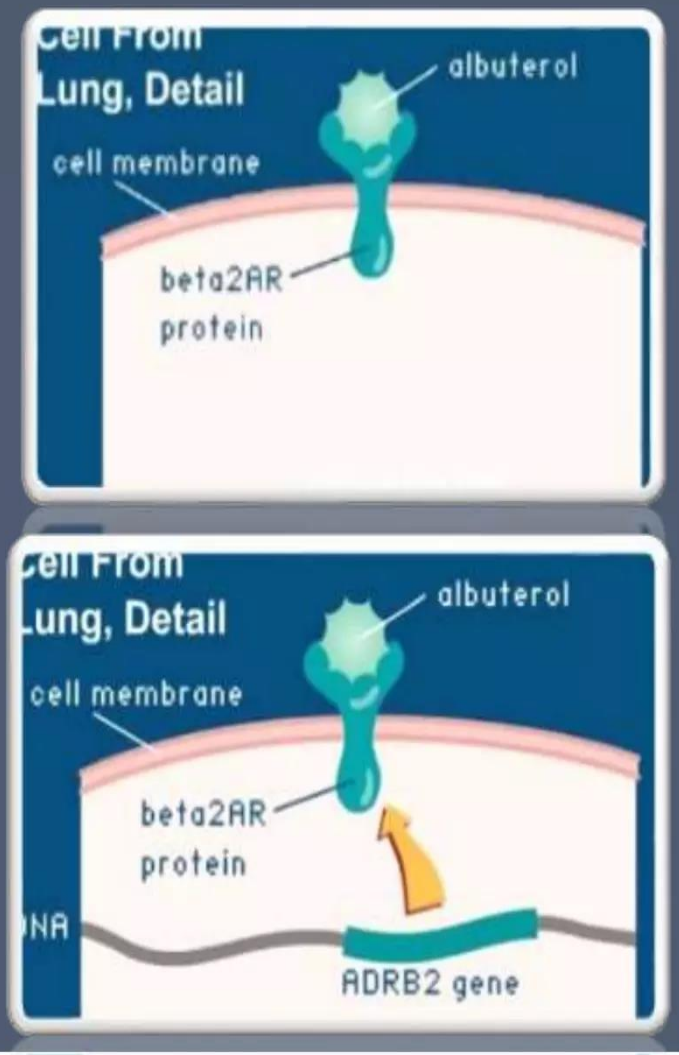
- ressecively inherited chronic hemolytic anemia
- nucleotide substition in b globin gene in chromosom 11 .

- high blood glucose
- insulin production is inadequate

Applying SNP profile to drug choices

- The drug albuterol is Commonly prescribed to relief symptoms of asthama
- albuterol effectively releife asthama but in some people not all. scientist are studying How people with different snp Response to treatment

- Albuterol acts on b-2-Adrenergic Receptor (beta 2 AR protein)
- Beta 2 AR encoded by AdRb2 gene



Conclusion

- Each detection method has its advantages and disadvantages.
- According to different research directions and purposes, researchers can choose suitable SNP detection methods. How to evaluate SNP detection methods depends on the following three aspects:
  - Design flexibility and success rate
  - Allele invocation and allele identification efficiency
  - Ease of running experiments and cost per sample