CHANDRA SHEKHAR AZAD UNIVERSITY OF AGRICULTURE AND TECHNOLOGY, KANPUR- 208 002

B.Sc. Horticulture Semester III
Course : Elementary Plant Biotechnology (BSH-211)

Lecture by: Sanjana Pathak

Teaching associate

Department of Genetics and Plant Breeding

Mutations in genomic DNA: cause polymorphism at locus among genotypes and make basis for markers

Base substitution

GATCCGAGTATCGCAATTAGCA
GATCCGAGTGTCGCAATTAGCA

Deletion

GATCCGAGTATCGCAATTAGCA
GATCCGAGTAATTAGCA

Insertion

GATCCGAGTATCGCAATTAGCA
GATCCGAGTATCGCAGCATTAGCA

Duplication

GATCCGAGTATCGCAATTAGCA
GATCCGAGTATCTCGCAATTAGCA

Inversion

GATCCGAGTATCGCAATTAGCA
GATGCCAGTATCGCAATTAGCA

Through long **evolutionary** accumulation, many different instances of mutation as mentioned above should exist in any given species

The **number** and **degree** of the **various types of mutations** define the **genetic diversity** within a species

This variation is useful, if it is **heritable** and **discernable**; as recognizable phenotypic variation or as genetic mutation distinguishable through **molecular marker technologies**

Molecular techniques to reveal genetic variation

Polymerase chain reaction (PCR)

Electrophoresis

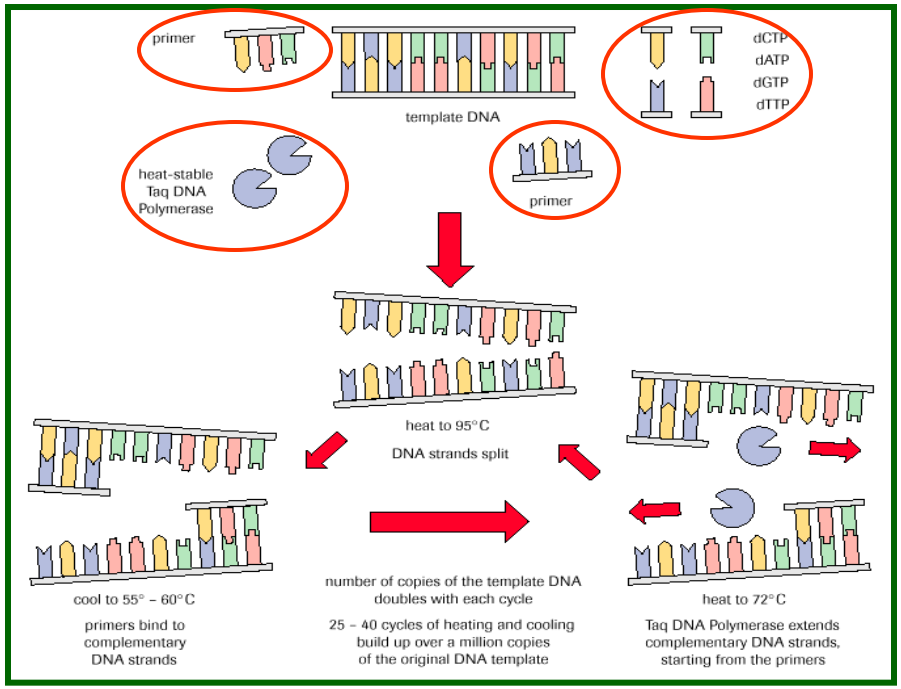
Hybridization

DNA sequencing

POLYMERASE CHAIN REACTION

PCR is a procedure used to amplify (make multiple copies of) a specific sequence of DNA

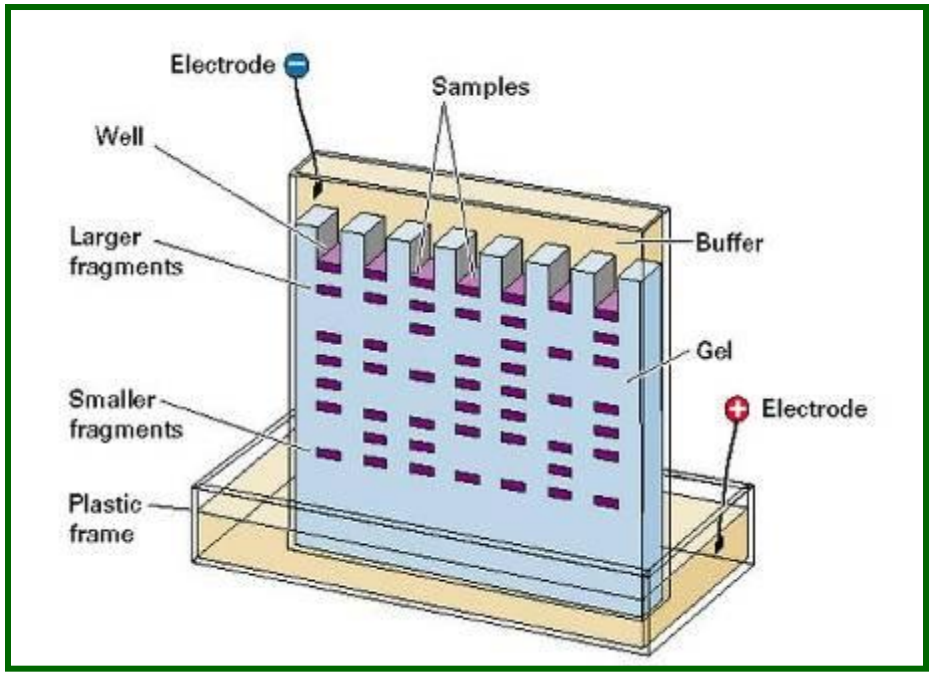
The method was invented by Kary Banks Mullis in 1983, for which he received the Nobel Prize in Chemistry ten years later



ELECTROPHORESIS

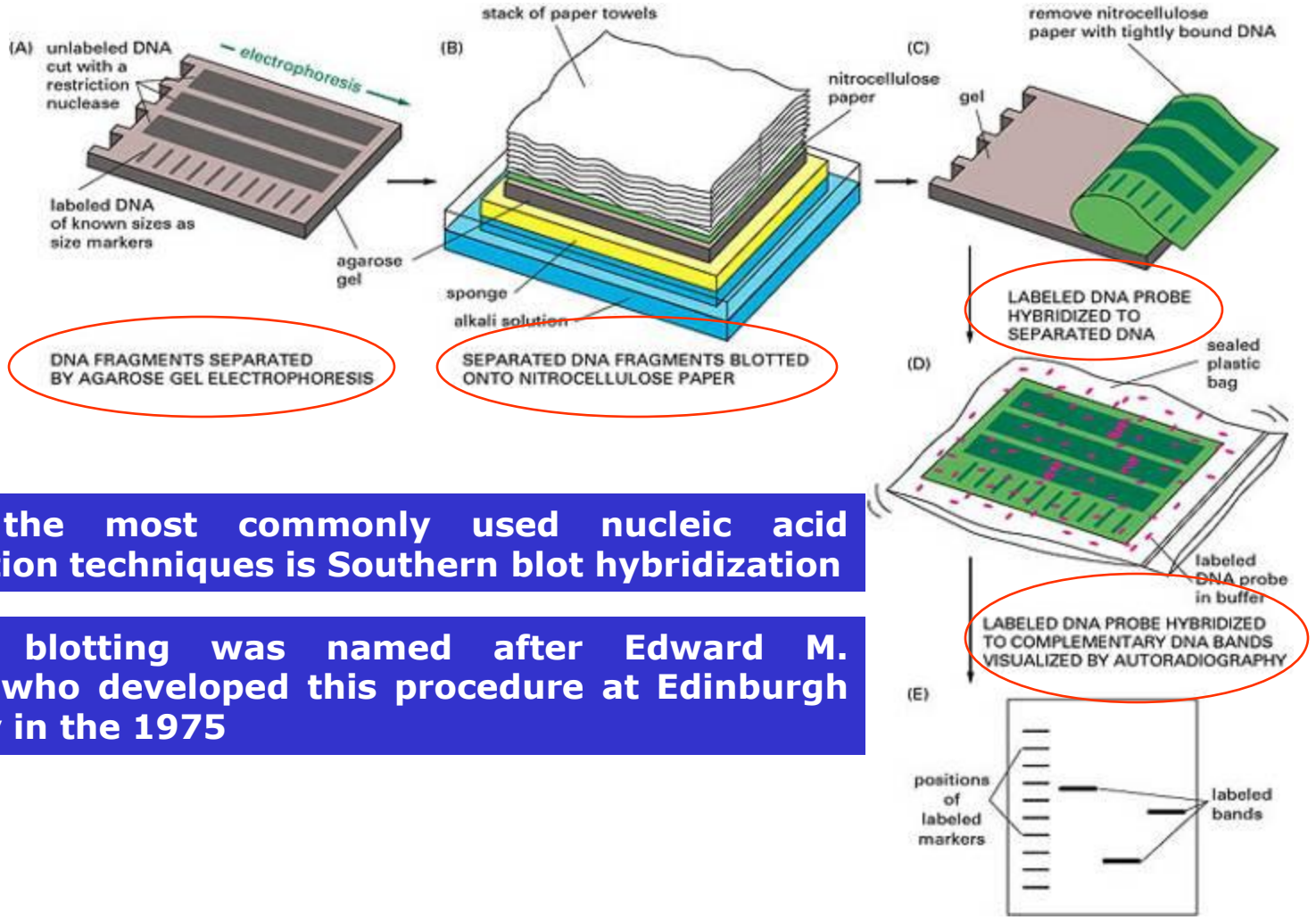
The term *'electrophoresis'* literally means "to carry with electricity"

Technique for separating the components of a mixture of charged molecules (proteins, DNAs, or RNAs) in an electric field within a gel or other support



Migration rate depend on electrical charge and size

HYBRIDIZATION



One of the most commonly used nucleic acid hybridization techniques is Southern blot hybridization

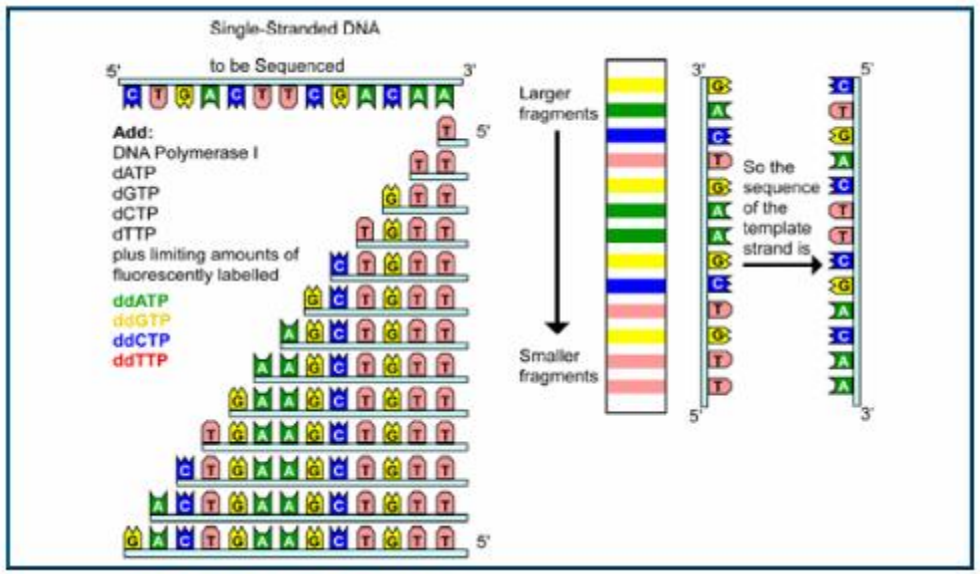
Southern blotting was named after Edward M. Southern who developed this procedure at Edinburgh University in the 1975

SEQUENCING

The process of determining the order of the nucleotide bases along a DNA strand is called sequencing

In 1977, 24 years after the discovery of the structure of DNA, two separate methods for sequencing DNA were developed: chain termination method and chemical degradation method

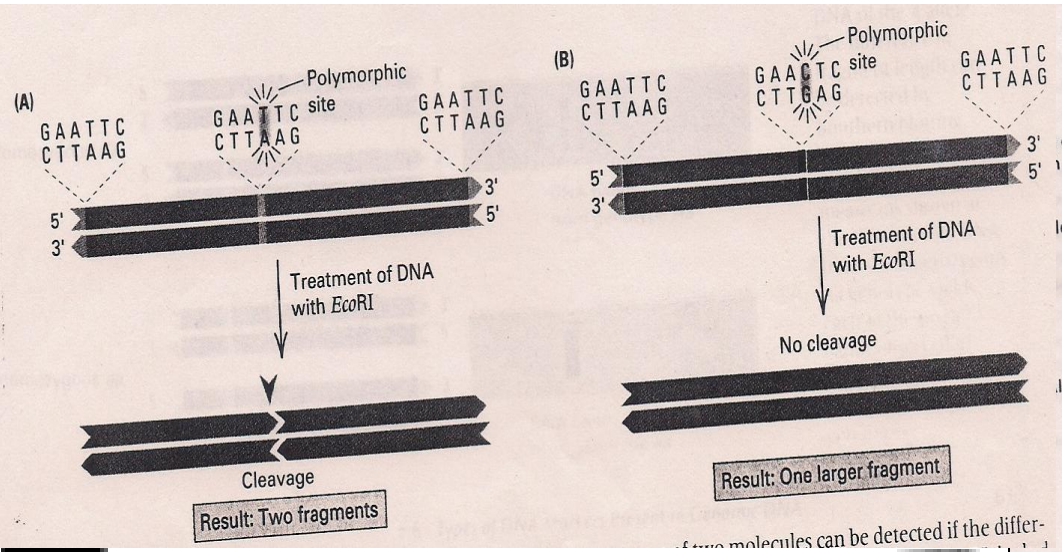
Principle: single-stranded DNA molecules that differ in length by just a single nucleotide can be separated from one another using PAGE

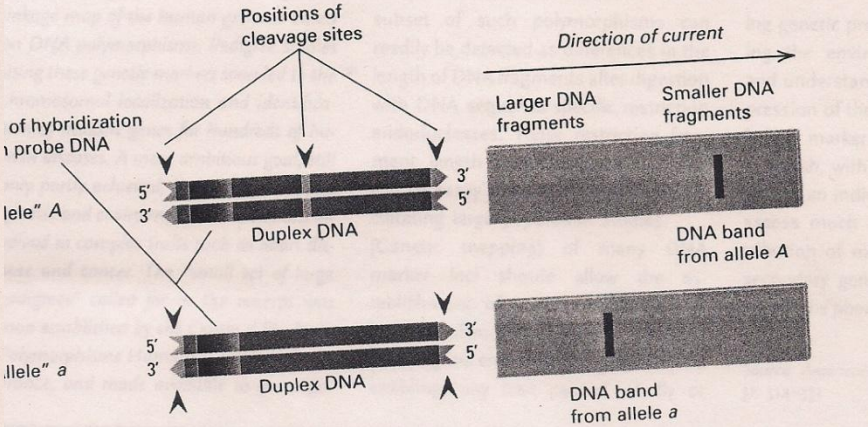


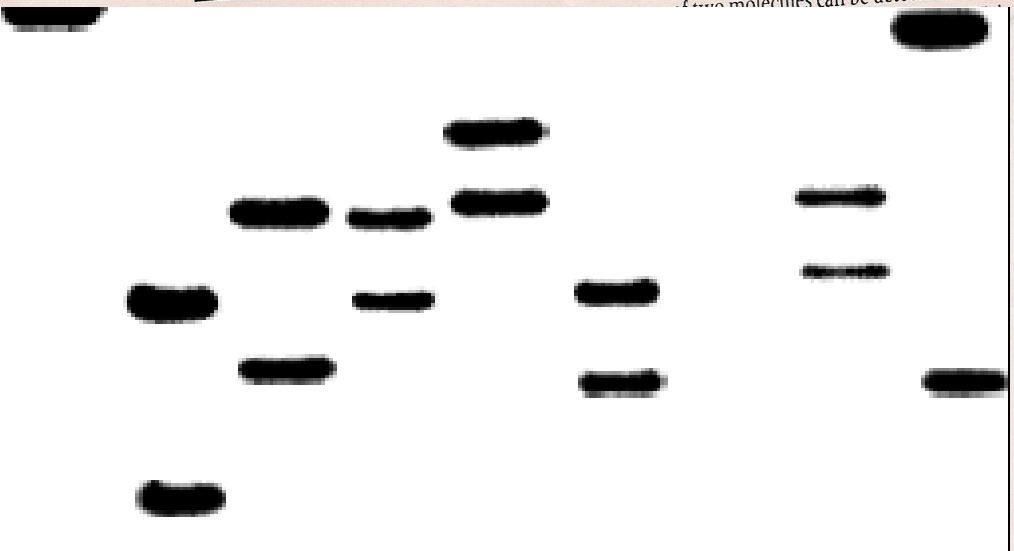
Chain elongation proceeds until, by chance, DNA polymerase inserts a dideoxynucleotide, blocking further elongation

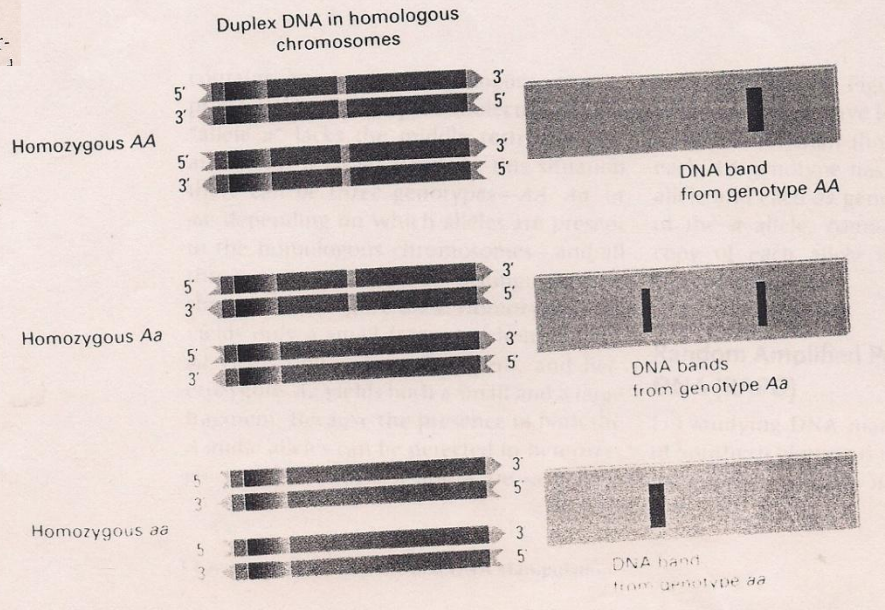
RFLP (Non-PCR based marker)

Targets variation in DNA restriction sites and in DNA restriction fragments. Sequence variation affecting the occurrence (absence or presence) of endonuclease recognition sites is considered to be main cause of length polymorphisms



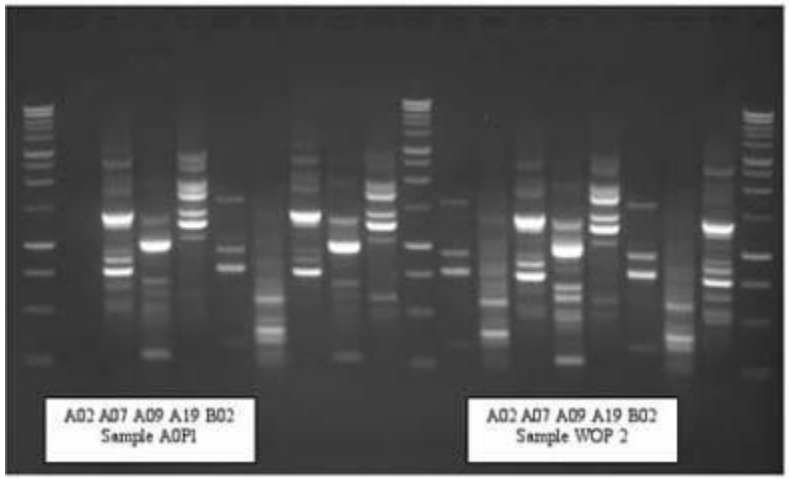


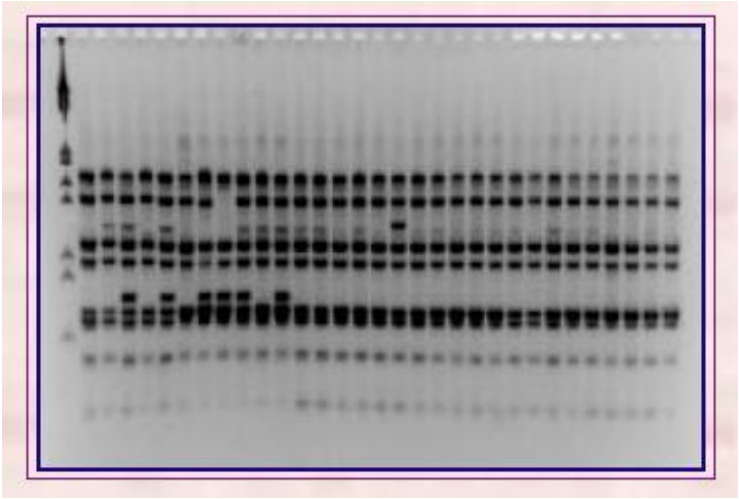




RAPD (PCR-based marker)

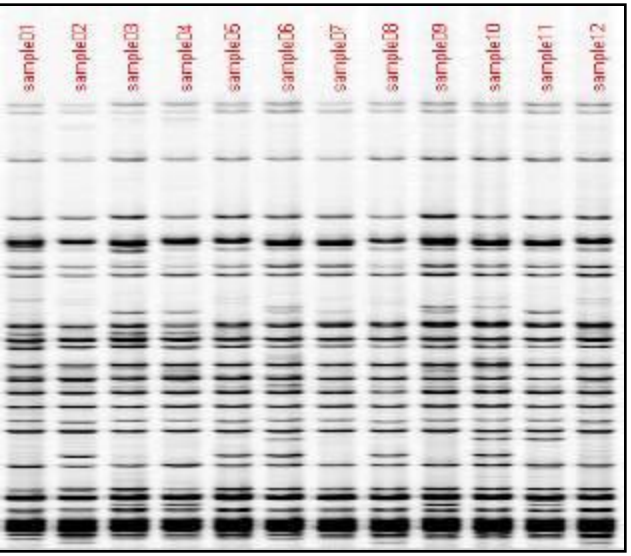
Uses primers of random sequence to amplify DNA fragments by PCR. Polymorphisms are considered to be primarily due to variation in the primer annealing sites, but they can also be generated by length differences in the amplified sequence between primer annealing sites

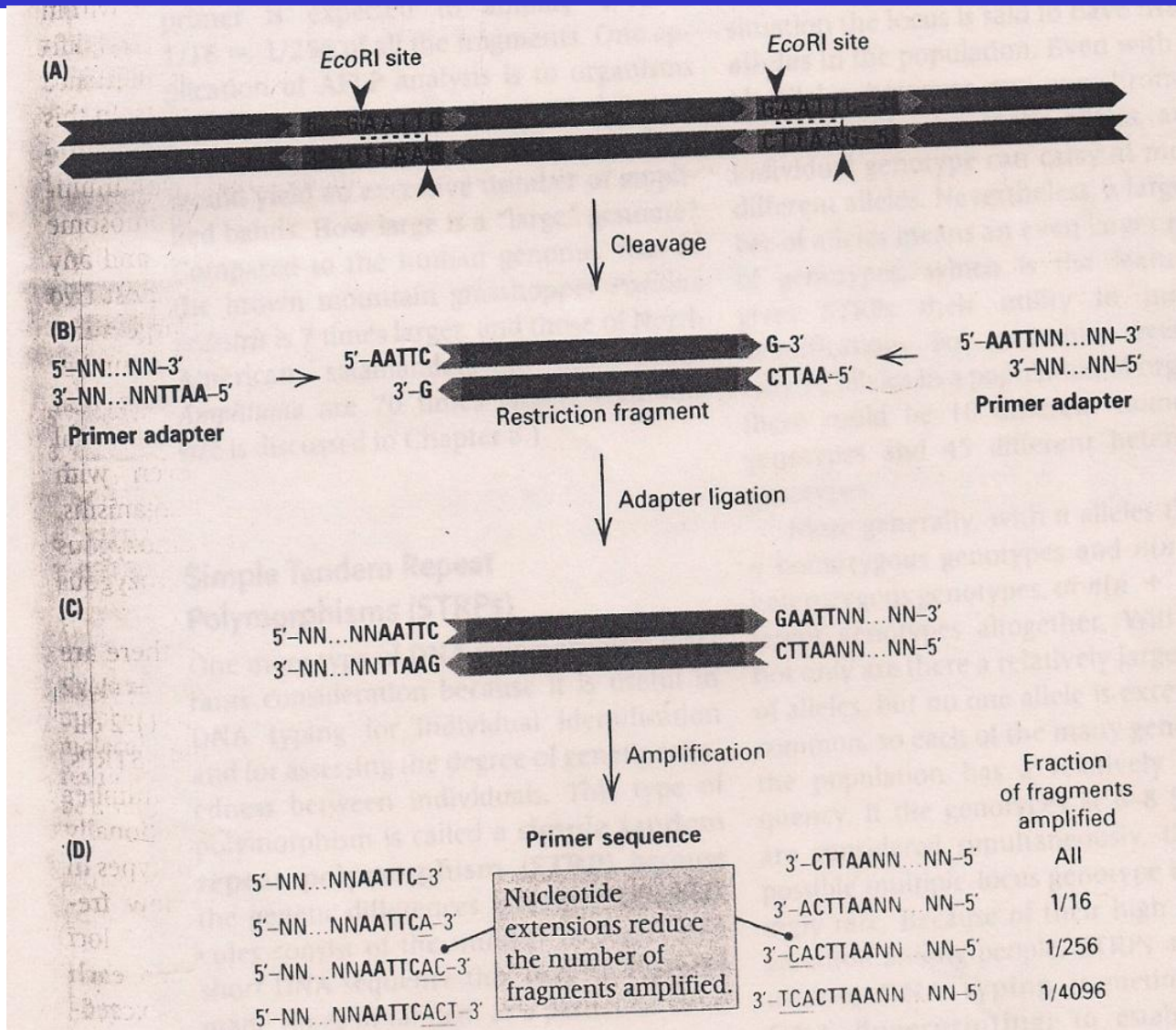




AFLP (PCR-based marker)

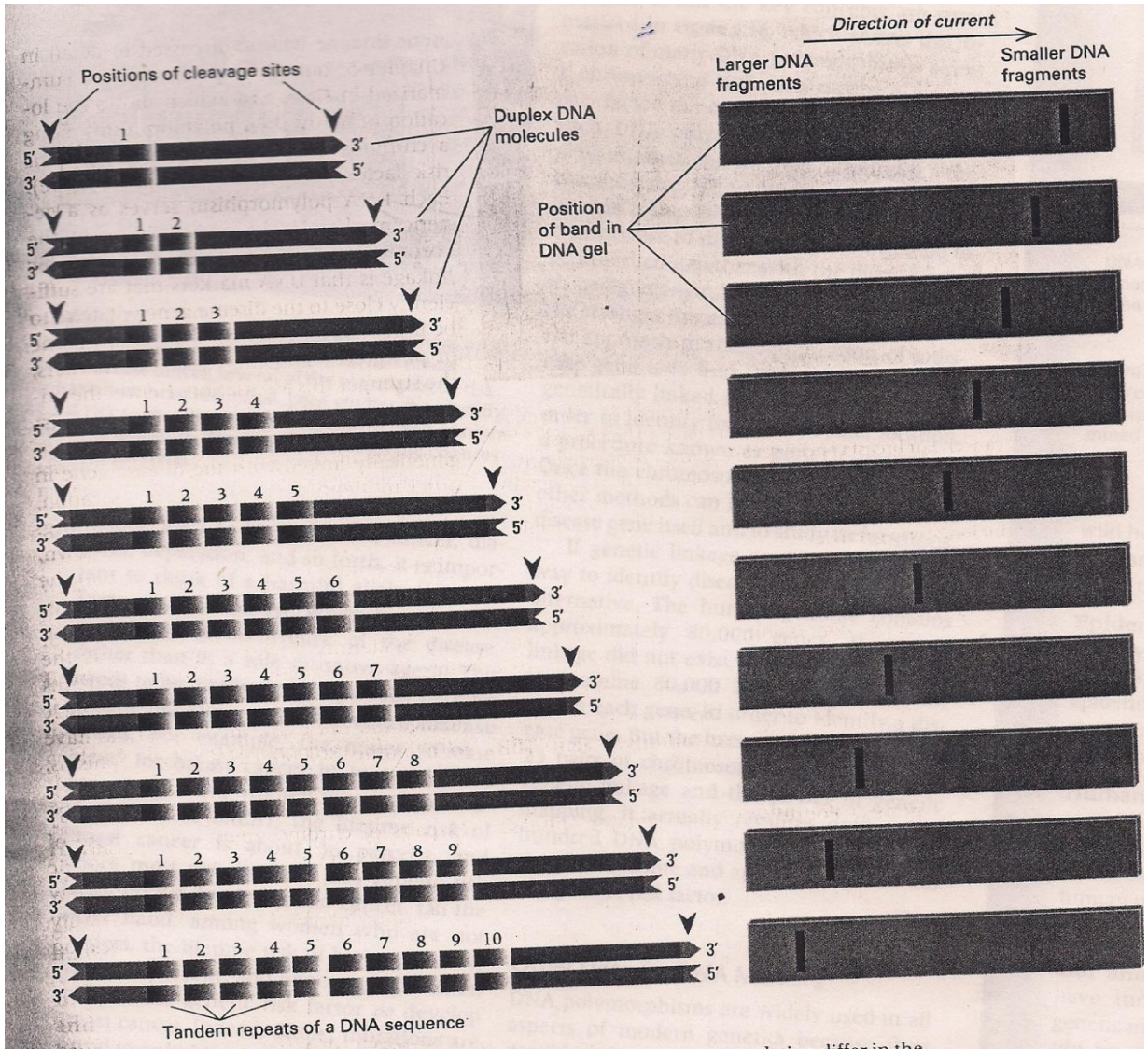
A variant of RAPD. Following restriction enzyme digestion of DNA, a subset of DNA fragments is selected for PCR amplification and visualization





Microsatellite (PCR based marker)

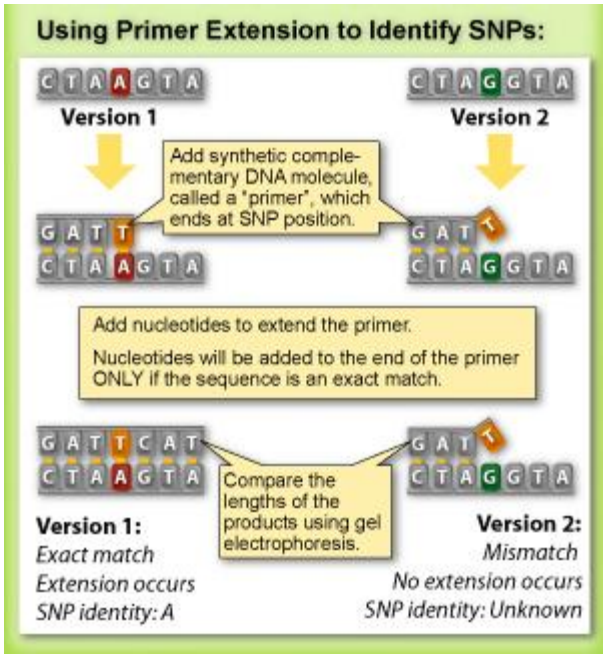
Targets tandem repeats of a small (1-6 base pairs) nucleotide repeat motif. Polymorphism due to the number of tandem repeats

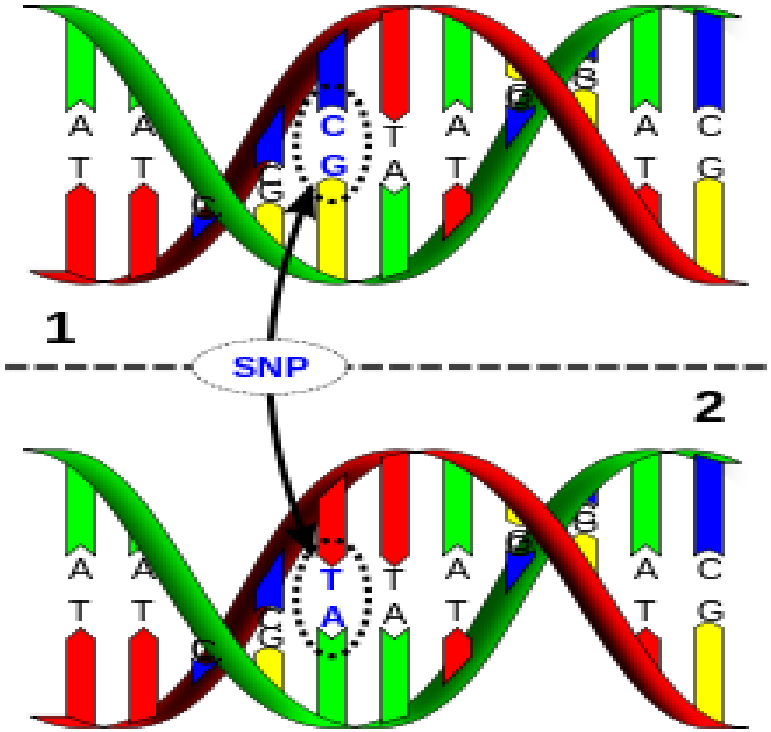


Single nucleotide polymorpphism

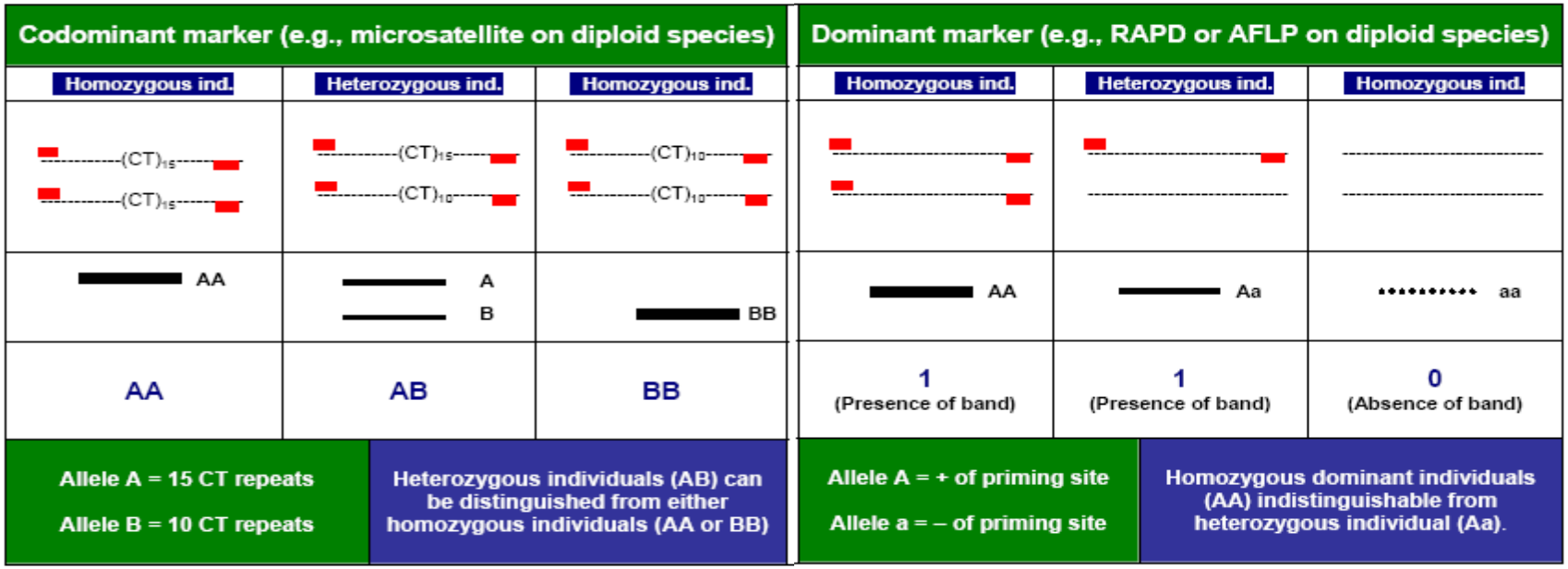
- Single nucleotide polymorphisms, frequently called SNPs (pronounced “snips”), are the most common type of genetic variation. Each SNP represents a difference in a single DNA building block, called a nucleotide. For example, a SNP may replace the nucleotide cytosine (C) with the nucleotide thymine (T) in a certain stretch of DNA.

- SNPs occur normally throughout DNA. They occur once in every 300 nucleotides on average, which means there are roughly 10 million SNPs in the human genome. Most commonly, these variations are found in the DNA between genes. They can act as biological markers, helping scientists locate genes that are associated with disease. When SNPs occur within a gene or in a regulatory region near a gene, they may play a more direct role in disease by affecting the gene’s function.





Codominance or dominace



Codominant marker:
A marker in which both alleles are expressed, thus heterozygous individuals can be distinguished from either homozygous state

Dominant marker:
A marker shows dominant inheritance with homozygous dominant individuals indistinguishable from heterozygous individuals

Categories of DNA based markers

- Based on the location, genomic region from where a marker is derived and functional properties of the marker region, DNA-based markers can be broadly divided into two broad categories:

- **Random DNA Markers (RDMs)**: Majority of the markers derived and used in the past were from the genomic DNA, and therefore could belong to either the transcribed or the non-transcribed part of the genome. The molecular markers derived from anonymous regions of the genome are called random DNA markers (RDMs), which may or may not be developed from the polymorphic site in gene or may not be developed from a gene.

- **Genic Molecular Markers (GMMs):** GMMs are developed from coding sequences like expressed sequence tags (ESTs) or from fully characterized genes and frequently such markers have been assigned known functions. Based on the site of polymorphism and effects on phenotypic variation, GMMs can be further divided into gene targeted markers and functional markers:

(i) Gene-targeted markers (GTMs): GTMs are those markers which have been derived from polymorphisms within genes, however not necessarily involved in phenotypic trait variation. For example, it may be based on the polymorphisms in un-translated regions (UTRs) of EST sequences.

(ii) Functional markers (FMs): FMs are derived from polymorphic sequences or sites within genes and, thus, more likely to be causally involved in phenotypic trait variation (e.g. candidate gene-based molecular markers). The FMs, depending on the involvement in the phenotypic trait variation, are further classified into two subgroups: (a) indirect functional markers (IFMs), for which the role for phenotypic trait variation is indirectly known, and (b) direct functional markers (DFMs), for which the role for the phenotypic trait variation is well proven.

Relevance in Plant Breeding

- Effective utilization of genetic resources through precise characterization, quantification and identification of genetic variation from all available sources. DNA markers are now seems to be an indispensable tool for characterizing genetic resources and providing breeders with more detailed information to assist in selecting parents.
- Tagging, cloning, and introgressing genes and/or quantitative trait loci (QTL) useful for enhancing the target trait using marker technologies. Markers have significant role in identifying rare recombinants with minimum linkage drag and recovering recurrent parent genotypes in back cross breeding programme.
- Manipulating (differentiating, selecting, pyramiding, and integrating) genetic variation in breeding populations.
- Markers techniques can also help plant breeding programs through assisting plant variety protection as well as in distinctness, uniformity, and stability (DUS) testing processes.
- Assessment of purity of cultivar especially markers can help to ensure the purity of hybrid cultivars.
- DNA markers have been and can be used to define heterotic groups that can be used to exploit heterosis (hybrid vigour). It is also possible to predict the heterosis using molecular markers, thus, time and resources involved in generating more number of crosses and further their evaluation could be minimised. It is true that molecular markers based prediction of heterosis is not conclusive so far, yet it is expected that further investigation particularly in functional markers system may help accurate prediction of heterosis in future.

Justification of markers use

- Early, quick and precise selection of a trait that is difficult to phenotype or has complex/low inheritance;
- Traits whose selection depends on specific environments or developmental stages that influence the expression of the target phenotype;
- Maintenance of the specific donor trait in backcross breeding with minimum linkage drag from the donor parent;
- Where recessive alleles are desirable and require to be restored during backcrossing or for speeding up backcross breeding in general
- Pyramiding of various genes which are difficult to phenotype in presence of the other trait;
- Selection of appropriate parents with diverse alleles.

Basic requirements

- Availability of a genetic map with an adequate number of uniformly-spaced polymorphic markers to accurately locate desired QTLs or major gene(s).
- There must be a close linkage between the QTL or a major gene of interest and adjacent markers. The success of MAS by and large depends on the localization of marker with respect to the target gene. The most desirable situation for MAS is when the molecular marker is located within the gene of interest. The localization of microsatellite markers within the opaque2 gene is an excellent example of this situation. Another situation encountered in breeding populations where markers are in linkage disequilibrium (LD) with the gene of interest throughout the population. LD is the tendency of certain combination of alleles to be inherited together. Selection using these markers can be called as LD-MAS. In the third situation, the marker is in linkage equilibrium (LE) with the gene of interest throughout the population. LE is in fact a random association of alleles at different loci and equals the product of allele frequencies within haplotypes, meaning that at random combination of alleles at each locus its haplotypes (combination of alleles) frequency has equal value in a population. This is the most difficult and challenging situation for applying MAS.
- There must be adequate recombination between the markers associated to the trait(s) of interest and the rest of the genome.
- Adequate infrastructural facilities particularly for high throughput genotyping and manpower support to analyze a large number of plants in a time and cost-effective manner.

Mapping population

- A population used for gene mapping is commonly called a mapping population. Mapping populations are usually obtained from controlled crosses.
- Decisions on selection of parents and mating design for development of mapping population and the type of markers used depend upon the objectives of experiments, availability of markers and the molecular map.
- The parents of mapping populations must have sufficient variation for the traits of interest at both the DNA sequence and the phenotype level. The variation at DNA level is essential to trace the recombination events.
- The more DNA sequence variation exists, the easier it is to find polymorphic informative makers.
- Selection of parents for developing mapping population is critical to successful map construction.
- Since a map's economic significance will depend upon marker-trait association, as many qualitatively inherited morphological traits as possible should be included in the genetic stocks chosen as parents for generating mapping population.