**Authors: Vivek Sharma and Ram Kunwar**

## Molecular markers types and applications

- A genetic marker is a gene or known DNA sequence on a chromosome that can be used to identify individuals or species.

- **Why we need Molecular Markers**

There will be no need if indentified traits have these three features

- Traits were easily score
- Individuals were easily classified into few distinct phenotypic classes
- Complete corresponding between phenotypic and genotypes

Based on Hybridization Markers are classified

1) Morphological Markers :- Height, Colour, Shape etc.
2) Biochemical Markers :- Isozyme Protein Banding Pattern
3) Molecular Markers

On the bases of chronology
1) First generation (RFLP and RAPD and there modifications)
2) Second generation (SSRs and AFLP and there modifications)
3) Third generation (ESTs and SNPs) markers.

Depending on the use of PCR
(1)PCR-based and
(2) non-PCR-based markers

Based on their molecular basis (Gupta et al. 2001).
(1) SNPs (generated by variation in DNA sequence)
(2) non-SNPs (produced by variation in sequence length, e.g., SSRs

On this basis of the location and the functional significance of markers
(1)Random, (2) gene-based, and (3) functional markers.

On the basis of the above and the throughput criteria,

(1) low-throughput hybridization-based markers,
(2) medium-throughput PCR-based markers,
(3) high-throughput sequence-based markers.

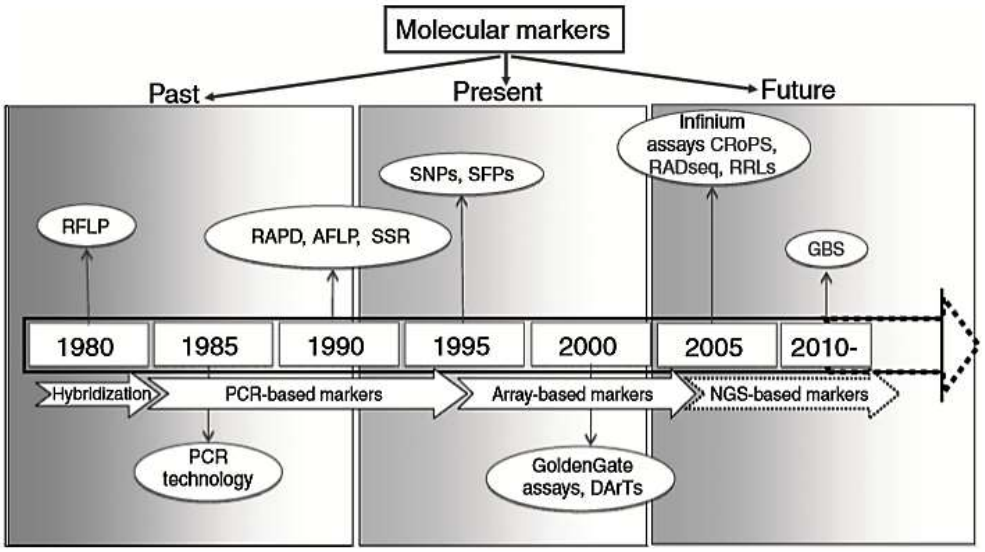



Image source: Mir et al., 2013

Molecular Markers variations are arise due to:

• Base pair changes.
• Rearrangements (translocation or inversion).
• Insertions or deletions.
• Variation in the number of tandem repeats.

Reflect heritable differences in homologous DNA sequences among individuals.

Advantage over previously detected makers

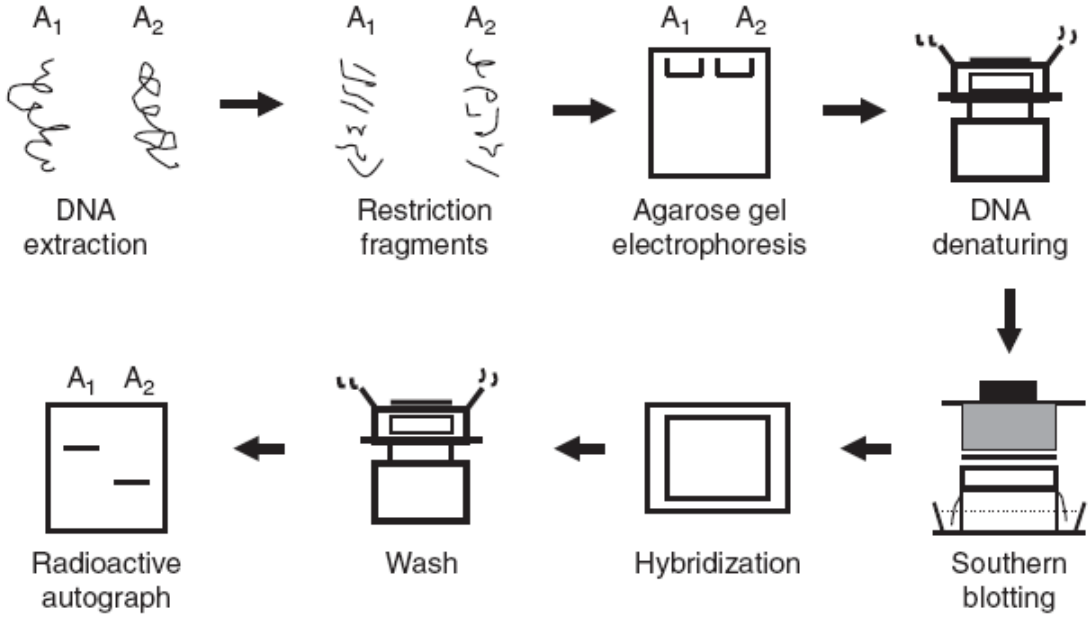
- Ubiquitous nature
- Stably inherited
- Multiple alleles for each marker
- Devoid of pleiotropic effects
- Detectable in all tissues, at all ages specially in early age to save time and cost, labour
- Long shelf life of the DNA samples.

Hybridization based (non-PCR) Technique: RFLPs **(Restriction Fragment Length Polymorphism analysis)** Botstein et al. (1980)

These genetic markers resulting from the variation or change in the length of defined DNA fragments produced by digestion of the DNA sample with restriction endonucleases and Electrophoretic comparison of the size of defined restriction fragments derived from genomic DNA .

Steps involved in

1. Isolate high quality DNA 2. Digest with a combination of restriction enzymes 3. Fractionate digested samples by electrophoresis 4. Transfer fragments to membrane 5. Hybridize with radioactively labeled DNA probe(s); detect by autoradiography or can also use non-radioactive labeling systems




Variation of RFLP analysis Polymorphism revealed by different probe/enzyme combinations among different accessions.

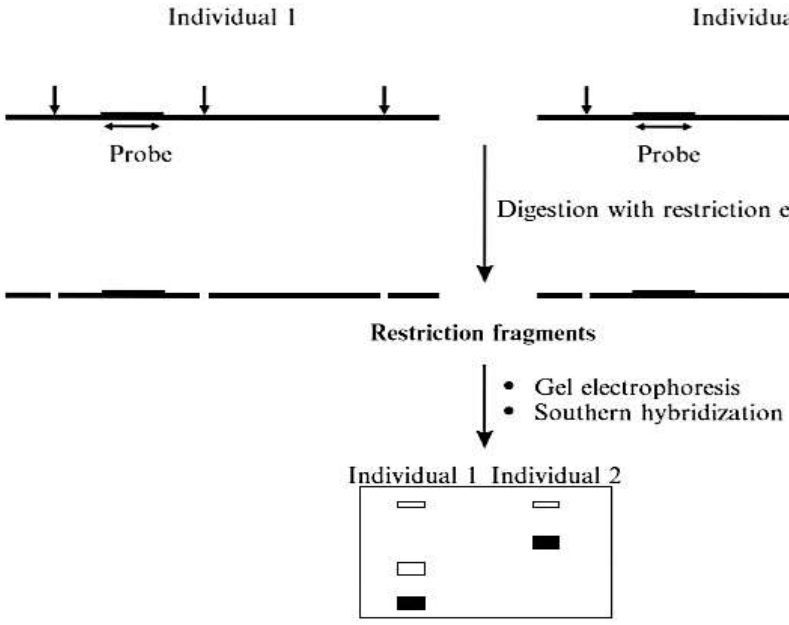


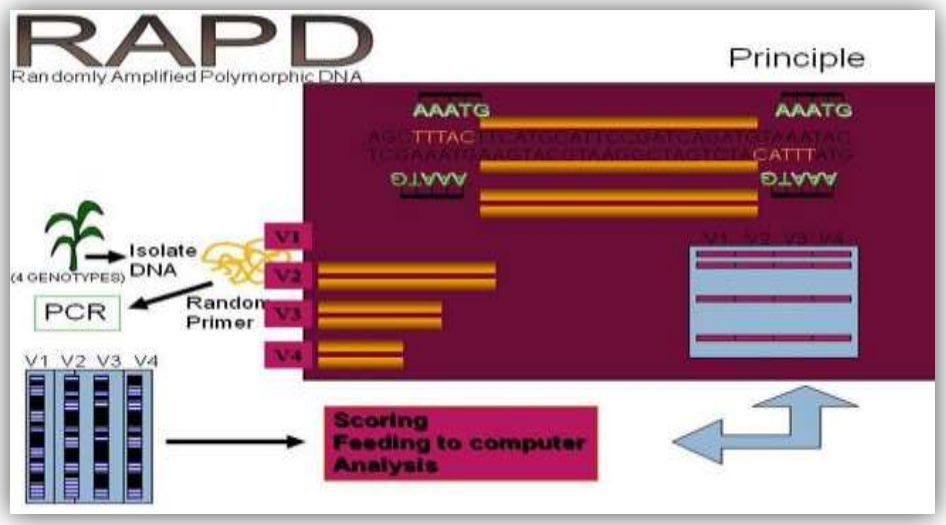

Image Source: Singh et al., 2015

Considerations for use of RFLPs - Relatively slow process -Use of radioisotopes has limited RFLP use to certified laboratories (but non-radioactive labelling systems are now in wide use) - Co-dominant markers; often species-specific - Need high quality DNA - Need to develop polymorphic probes – expensive

**PCR based techniques (RAPD, ISSR, SSR, AFLP, EST , SNP)**

**Random Amplified Polymorphic DNA (RAPD)** are genetic markers resulting from PCR amplification of

genomic DNA sequences recognized by random primers of arbitrary nucleotide sequence (Williams et al., 1990). RAPDs are dominant markers that require no prior knowledge of the DNA sequence, which makes them very suitable for investigation of species that are not well known (Williams et al. 1993).




A single, random-sequence oligonucleotide primer in a low stringency PCR (35–45°C) simultaneously Amplifies several discrete DNA fragments

random amplified polymorphic DNA (RAPD) by Williams *et al.* (1990)
arbitrary primed PCR (AP-PCR) by Welsh and McClelland (1990)
DNA amplification fingerprinting (DAF) by Caetano-Anollés *et al*., (1991)

10-mer oligonucleotide
several discrete DNA products up to 3 kb are amplified (amplicons)
these are considered to originate from different genetic loci
visible in conventional agarose gel electrophoresis as the presence or absence of a particular RAPD band
RAPDs predominantly provide dominant markers

Application

- Construction of genetic map eg Arabidopsis
- Mapping of traits
- Fingerprinting of individuals
- Identification of somatic hybrids

**AFLP (Amplified Fragment Length Polymorphisms)** is a combination of PCR and RFLP Involved in Informative DNA fingerprints of amplified fragments. It is based on the selective amplification of a subset of genomic restriction fragments using PCR.
Steps involved in AFLP process 1. Digest genomic DNA with restriction enzymes 2. Ligate with adaptors (defined sequences) to both ends of the fragments 3. Carry out PCR on the adaptor-ligated mixture, using primers that target the adaptor, but that vary in the base(s) at the 3' end of the primer. AFLP technology is a DNA fingerprinting technique that combines RFLP and PCR.
Advantages of AFLP's

- Very sensitive
- Good reproducibility but technically demanding

- Relatively expensive technology
- Discriminating homozygotes from heterozygotes
- Requires band quantitation (comparison of pixel density in images from a gel scanner)

**SSR (Single Sequence repeats) variants**

1. Microsatellites
2. Short tandem repeats (STRs)
3. Sequence-tagged microsatellite sites (STMS)

- Tandemly arranged repeat units 1–6 bp long
- Di-, tri- and tetranucleotide repeats – (CA)n, (AAT)n and (GATA)n with different length of reaptes motif formed by slipped strand mispairing
- C. widely distributed in genomes (plants &animals (Tautz and Renz, 1984).

Construction steps

I) Microsatellite library construction
II) Identification of unique microsatellite loci
III) Identification a suitable area for primer design
IV) Obtaining a PCR product
V) Evaluation of banding pattern
VI) Assessing PCR product for polymorphisms

Advantages of SSRs:

- Hypervariability, Reproducibility, Codominant nature
- Locus specificity random dispersion throughout most genomes
- More variable than RFLPs or RAPDs.
- Readily analysed by PCR and easily detected on PAGE SSLPs with large size differences - detected on agarose gels SSR markers can be multiplexed genotyping throughput is high and can be automated start-up costs are low for manual assay methods (once the markers have been developed) SSR assays require only very small DNA samples(ca.100 ng / individual)

The disadvantages

- Labour intensive development process
- High start-up costs for automated methods.

**Inter-Simple Sequence Repeats (ISSR)**

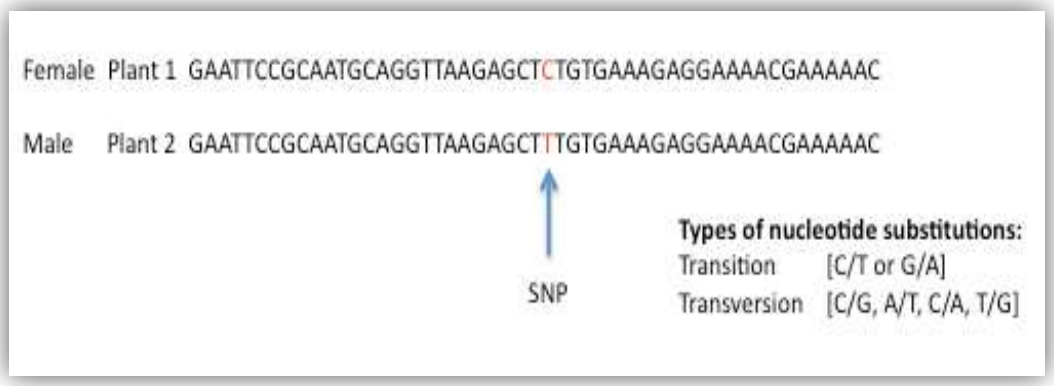
- Amplification of DNA samples present between two SSRs opposite oriented direction
- SSR used as primers for IISR

Advantage: Highly polymorphic, No sequence required
Limitation: Dominant, Reproducibility, Homology same as RAPD

**SNP (Single Nucleotide Polymorphisms)** Pronounced as snip - an individual nucleotide base difference
There are three types recognized

- Transitions (C/T or G/A)
- Transversions (C/G, A/T, C/A or T/G) e.g., AAGCCTA AAGCTTA The two alleles are C and T.
- Insertions and deletions (Indels)

Some of more interesting point about sinps




- Human genome has at least 1.42 million SNPs as comprising to 100000 of which result in an RFLP
- 1% of the population should have SNP 90% of all human genetic variation are SNPs and occur every 100–300 bases
- C/T transitions constitute 2/3 or 67% of the SNPs
- Single base variants in cDNA (mRNA) or single Nucleotide base - the smallest unit of inheritance,

Typical SNP frequencies are in the range of one SNP every 100–300 bp
SNPs may fall within coding sequences of genes –

- if same polypeptide then called synonymous SNP or Silent SNP its does not altered amino acid
- if different polypeptide then called non-synonymous i.e. altered amino acid sequence

Non-coding regions of genes may have SNPs by

- Gene splicing,
- Transcription factor binding the sequence of non-coding RNA
- The intergenic regions between genes at different frequencies in different chromosome regions

Approaches adopted for discovery of novel SNPs:
I) in vitro discovery, where new sequence data is generated
II) in silico methods that rely on the analysis of available sequence data
III) Indirect discovery, where the base sequence of the polymorphism remains unknown SNP

Sobrino et al. (2005) classified SNP genotyping assays into 4 groups (based on the molecular mechanisms)

- Allele-specific hybridization
- Primer extension
- Oligonucleotide ligation
- Invasive cleavage

The advantages of SNPs are their most abundant in numbers even located within gene would provide ideal marker and do not involve gel electrophoresis. SNP detection is more rapid because it is based on oligonucleotide hybridization analysis.

**Application of Molecular markers**

- Phylogenetic studies
- Trait Identification and Mapping
- DNA finger printing
- Genetic diagnostics
- Expression Profile Analysis
- Study of genome
- Gene mapping / Gene tagging
- Seed testing
- Identifying location of QTL's
- Marker Assisted Selection (MAS)
- Marker Assisted Backcrossing Breeding (MABB)

**References (if any)**

1. B.D. Singh and A.K. Singh, Marker-Assisted Plant Breeding: Principles and Practices, DOI 10.1007/978-81-322-2316-0_2, # Author(s) 2015

2. Reyazul R. Mir and Rajeev K. Varshney; Molecular Markers in Plants, First Edition. Edited by Robert J. Henry. C 2013 John Wiley & Sons, Inc. Published 2013 by John Wiley & Sons, Inc.

## Terms - Do not remove or change this section ( It should be emailed back to us as is)