RAMA UNIVERSITY

www.ramauniversity.ac.in

FACULTY OF ENGINEERING &TECHNOLOGY
DEPARTMENT OF BIOTECHNOLOGY

Dr. Simranjit Singh
Assistant Professor
Department of Biotechnology
Rama University, Kanpur

DNA based Markers-RFLP, AFLP, RAPD, SSR, MAS

Simranjit Singh, Ph.D.

DNA markers

Appendix

- Molecular Markers
- Basic Principle of Molecular Markers Detection
- Molecular Markers based on DNA Hybridization
- Markers Based on PCR Amplification
- Random amplified polymorphic DNA (RAPD)
- Amplified fragment length polymorphism (AFLP)
- Molecular Marker Assisted Selection

Molecular Markers

- A **molecular marker** is a **DNA sequence** in the genome which can be located and identified.
- As a result of genetic alterations (**mutations, insertions, deletions**), the base composition at a particular location of the genome may be **different in different plants.**
- These differences, collectively called as **polymorphisms** can be mapped and identified.
- Plant breeders always prefer to **detect the gene** as the **molecular marker**, although this is not always possible.
- The alternative is to have **markers** which are closely associated with **genes** and inherited together.

Molecular Markers

The molecular markers are highly reliable and advantageous in plant breeding programmes:

i. Molecular markers provide a true representations of the **genetic makeup** at the **DNA level.**

ii. They are consistent and not affected by **environmental factors.**

iii. Molecular markers can be detected much before **development of plants** occur.

iv. A **large number of markers** can be generated as per the **needs.**

Basic Principle of Molecular Markers Detection

Let us assume that there are **two plants** of the **same species**—one with **disease sensitivity** and the other with **disease resistance**.

- If there is **DNA marker** that can identify these **two alleles**, then the **genome can be extracted**, **digested by restriction enzymes**, and **separated by gel electrophoresis.**
- The **DNA fragments** can be detected by **their separation**. For instance, the **disease resistant plant** may have a **shorter DNA fragment** while the **disease — sensitive plant** may have a **longer DNA fragment** (Fig. 53.1).

Molecular markers are of two types:

1. Based on nucleic acid **(DNA) hybridization** (non-PCR based approaches).
2. Based on **PCR amplification** (PCR-based approaches).

Molecular Markers

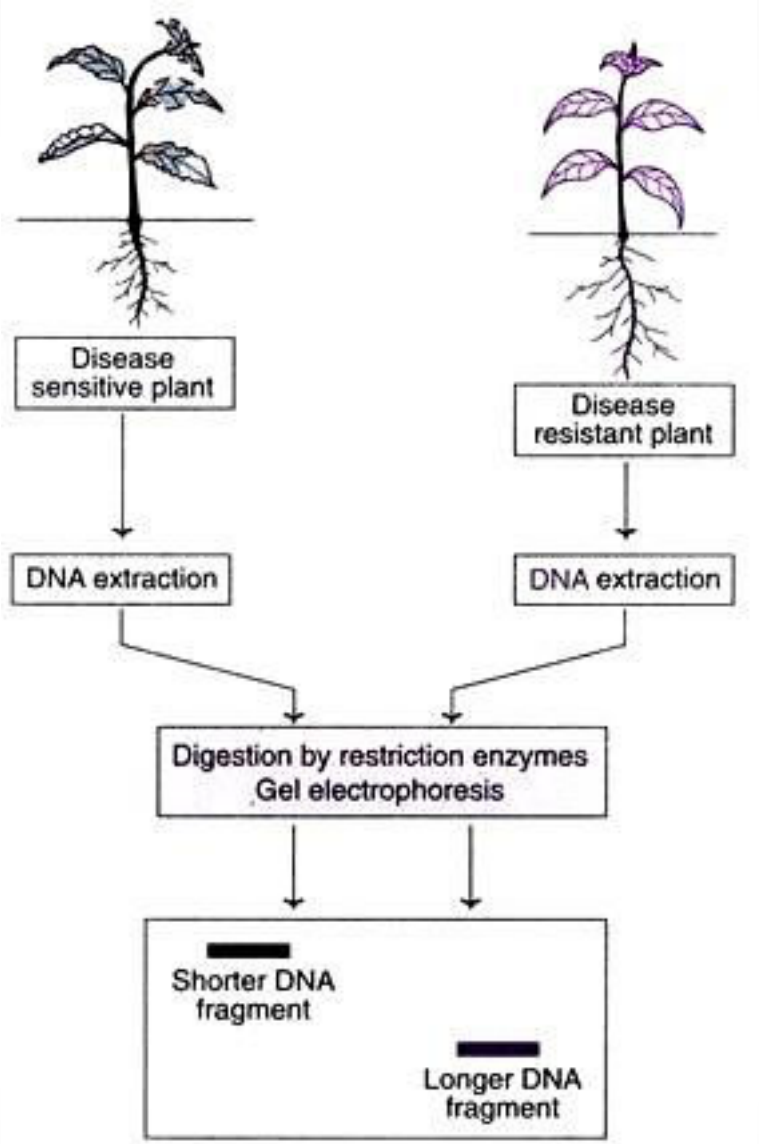


Fig. 53.1 : Basic principle of molecular marker detection (screening of genotypes for the identification of DNA markers).

Molecular Markers based on DNA Hybridization

The DNA piece can be cloned, and allowed to hybridize with the genomic DNA which can be detected. **Marker-based DNA hybridization** is widely used. The major **limitation** of this approach is that it requires **large quantities of DNA** and the **use of radioactivity** (labeled probes).

Restriction fragment length polymorphism (RFLP):

- RFLP was the very first technology employed for the detection of **polymorphism**, based on the **DNA sequence differences.**
- **RFLP** is mainly based on the **altered restriction enzyme sites**, as a result of **mutations** and **re-combinations** of **genomic DNA.**
- An outline of the RFLP analysis is given in **Fig. 53.2**, and schematically depicted in **Fig. 53.3**.
- The procedure basically involves the **isolation of genomic DNA**, its **digestion by restriction enzymes**, **separation by electrophoresis**, and finally **hybridization by incubating with cloned and labeled probes** (**Fig. 53.2**).

Molecular Markers based on DNA Hybridization

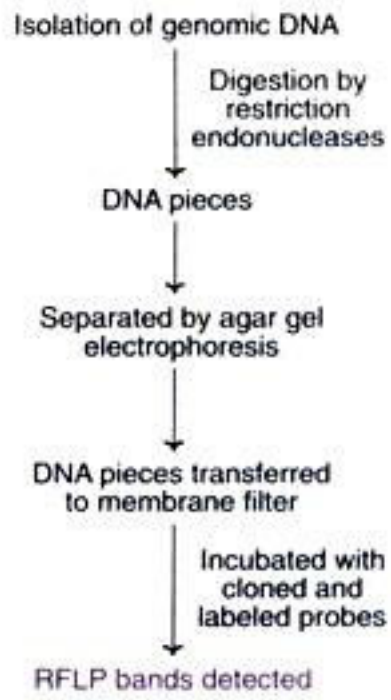


Fig. 53.2 : An outline of restriction fragment length polymorphism (RFLP) analysis as a molecular marker in plant breeding.

Molecular Markers based on DNA Hybridization

Based on the **presence of restriction sites**, **DNA fragments** of different **lengths** can be generated by using **different restriction enzymes**. In the Fig. 53.3, two DNA molecules from **two plants (A and B)** are shown. In **plant A**, a mutations has occurred leading to the loss of restriction site that can be digested by **EcoRI.**

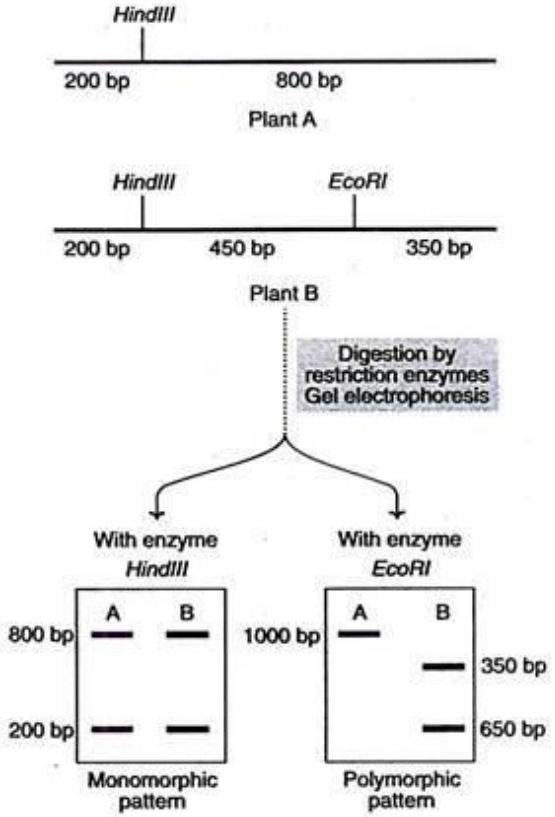
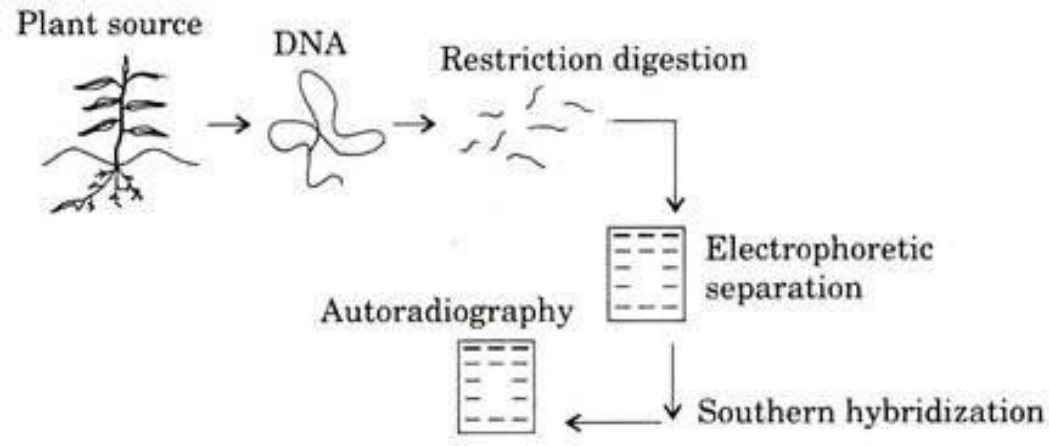


Fig. 53.3 : A schematic representation of restriction fragment length polymorphism (RFLP) analysis as molecular marker.

The result is that when the DNA molecules are digested by the enzyme **HindIII**, there is no difference in the DNA fragments separated. However, with the enzyme **EcoRI**, plant A DNA molecules is not digested while plant B DNA molecule is digested. **This results in a polymorphic pattern of separation.**

Molecular Markers based on DNA Hybridization

Based on the **presence of restriction sites**, **DNA fragments** of different **lengths** can be generated by using **different restriction enzymes**. In the Fig. 53.3, two DNA molecules from **two plants (A and B)** are shown. In **plant A**, a mutations has occurred leading to the loss of restriction site that can be digested by **EcoRI.**



The result is that when the DNA molecules are digested by the enzyme **HindIII**, there is no difference in the DNA fragments separated. However, with the enzyme **EcoRI**, plant A DNA molecules is not digested while plant B DNA molecule is digested. **This results in a polymorphic pattern of separation.**

Markers Based on PCR Amplification

Polymerase chain reaction (**PCR**) is a novel technique for the amplification of **selected regions of DNA** .The advantage with PCR is that even a minute quantity of DNA can be amplified. Thus, PCR-based molecular markers require only a small quantity of DNA to start with

The result is **PCR-based markers may be divided into two types:**

1. **Locus non-specific markers** e.g. **random amplified polymorphic DNA** (RAPD); **amplified fragment length polymorphism** (AFLP).
2. **Locus specific markers** e.g. **simple sequence repeats** (SSR); **single nucleotide polymorphism** (SNP).

Random amplified polymorphic DNA (RAPD)

Random amplified polymorphic DNA (RAPD) markers:

- RAPD is a **molecular marker** based on **PCR amplification**. An outline of RAPD is depicted in Fig. 53.4. The **DNA isolated** from the genome is **denatured** the template molecules are **annealed with primers**, and **amplified by PCR.**

- Single **short oligonucleotide primers** (usually a 10-base primer) can be arbitrarily selected and used for the amplification DNA segments of the genome (which may be in distributed throughout the genome). T**he amplified products are separated on electrophoresis and identified.**

- When the DNA molecules are digested by the **enzyme Hindlll**, there is no difference in the DNA fragments separated. However, with the enzyme **EcoRI**, plant A DNA molecules is not digested while plant B DNA molecule is digested. This results in a polymorphic pattern of separation.

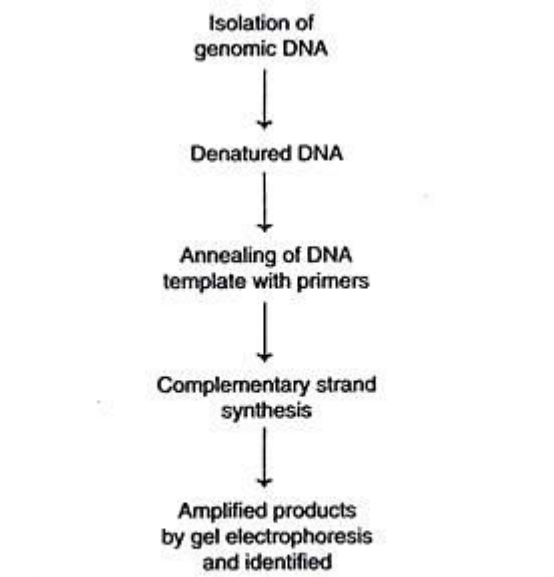


Fig. 53.4 : An outline of random amplified polymorphic DNA (RAPD) analysis as a molecular marker in plant breeding.

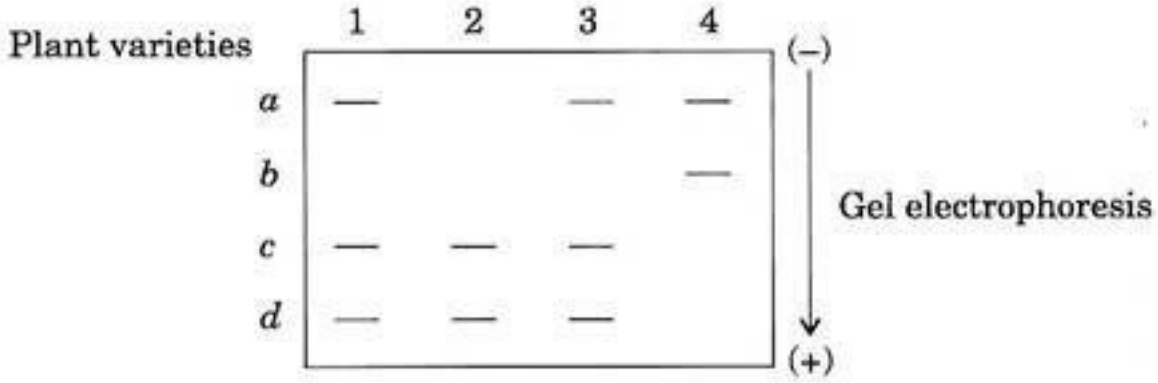
Random amplified polymorphic DNA (RAPD)

Applications of RAPD marker in plants.

1. RAPD is used to **distinguish between variety** is based on difference in **DNA sequence**. RAPD have been used to identify nearly **15 commercial sunflower varieties.** The new **bean varieties (Phaseolus vulgaris)** which are difficult to distinguish based on morphological trait have been used as ideal candidate for the application of **RAPD marker methods** when DNA was extracted from each varieties. The 12 samples were analysed using **60 primers.**

This produced **296 markers** and that could be scored. Almost 85% similarity was predicted. For example, DNA from plant allows the amplification of sequences **a, c, d** but **not b**. This indicates that in **plant 1**, primer sites for the primers used not found at **sequence b**. Similarly, a DNA sequence alternation at one of the primer binding site (priming) for a sequence 'a' has prevent it from being amplified when DNA from **plant 2 is used.**

Random amplified polymorphic DNA (RAPD)



RAPD have been extensively used for number of horticultural crops in variety identification, genetic purity and sex determination. A specific RAPD marker has been used to select for high and low β-glucan content between barley varieties.

Random amplified polymorphic DNA (RAPD)

2. RAPD markers are employed in the construction of genetic maps. Genetic maps of several plants including model plant Arabidopsis and tobacco have been constructed. RAPD markers have been used to construct 15 linkage groups in coffee. Both genomic and chloroplast DNA provided the source of probes.

RAPD markers are used for the selection of segregating populations more or less indirectly, during plant breeding. These markers also accelerate back crossing process and allow the selection of individually with more of recurrent genome at each generation facilitate breeding programme to be completed within few generations.

3. RAPD molecular marker used in the direct selection of desirable trait. Molecular marker linked to the trait of interest can be screened for at any stage in the breeding programme.

Random amplified polymorphic DNA (RAPD)

4. RAPD and other molecular markers have great value in the selection for desirable trait in long-lived species which takes long time for maturity and show phenotypic character. For example, avacado (Persea americana) fruit quality can be assessed in seedling itself using RAPD molecular marker.

5. **RAPD markers** have been used to identify **several disease resistant genes** in plants. The **rp94 gene** is responsible for **resistance** to **stem rust** (Puccinia gramnis) in **barley**. **RAPD markers** identified to link to this gene. Similarly, **RAPD markers** linked to **heat smut resistance** gene have been characterized. Controlling of height in barley plant by specific gene has been used to locate dwarfism gene by RAPD marker.

Random amplified polymorphic DNA (RAPD)

- 6. In tissue culture work, somatic hybrids involving protoplast fusion requires thorough screening. However, screening of somatic hybrid is cumbersome. Therefore, RAPD markers can be exploited in identifying somatic hybrids. RAPD analysis provides an important tool for the characterization of biodiversity.

- Identification of areas rich in endemic genotypes helps in habitat conservative and prevents species extinction. Molecular analysis of genetic diversity using RAPD or RFLP in plant genetic germplasm collection facilitate better management especially space and resources are serious constraints.

- RAPD analysis has been used for the identification of duplicates in germplasm. These duplicates are then discarded once no morphological differences were detected. RAPD analysis has been implicated in the analysis of rice genome collections held at the International Rice Research Institute, Phillipines.

- Genetic diversity was carried out in a set a 63 tetraploid wheat genotype. Which comprises 24 duran land races, 18 duran cultivars and nine diococcum cultivars, two wild tetraploid species? The duran and dicoccum wheat genotypes are part of the germplasm used in Indian tetraploid wheat breeding programme.

Random amplified polymorphic DNA (RAPD)

- RAPD scoring analysis reveals 78% were polymorphic in different categories of Indian tetraploid wheat. These indicate that RAPD diversity data can be used in breeding improved cultivars and maintaining genetic diversity in germplasm. Similarly RAPD markers were generated from 6 groups of 23 varieties of Tibetian barley. Nearly 23 RAPD and 29 genes loci were identified on 72 chromosomes.
- RAPD have also been used in variety identification and purity in grain processing in food industry. For example, particular duran variety of wheat is used in the preparation of food products. Contamination of other variety can be identified by using these molecular markers.

Demerits of RAPD:

1. Constraints about reproducibility of results.
2. Since RAPD markers are dominant, only half the genetic information are co-dominant markers.
3. Null alleles not directly detected.

Markers Based on PCR Amplification

Based on the **nucleotide alterations in the genome**, the polymorphisms of amplified DNA sequences **differ** which can be identified **as bends on gel electrophoresis.**

Genomic DNA from two different plants often results in different amplification patterns i.e. RAPDs.

This is based on the fact that a **particular fragment of DNA** may be generated from **one individual,** and not from others. This represents **polymorphism** and can be used as a **molecular marker** of a particular species.

Amplified fragment length polymorphism (AFLP)

- AFLP is a novel technique involving a **combination of RFLP and RAPD**. AFLP is based on the principle of **generation of DNA fragments using restriction enzymes** and **oligonucleotide adaptors** (or linkers), and their **amplification by PCR**. Thus, this technique combines the usefulness of **restriction digestion and PCR.**

- The **DNA** of the genome is **extracted.** It is subjected to **restriction digestion** by **two enzymes** (**a rare cutter e.g. MseI**; **a frequent cutter e.g. EcoRI**). The cut ends on both sides are then ligated to known sequences of **oligonucleotides** (Fig. 53.5).

Amplified fragment length polymorphism (AFLP)

- **PCR** is now performed for the **pre-selection of a fragment of DNA** which has a single specific nucleotide. By this approach of pre-selective amplification, the pool of fragments can be reduced from the original mixture. In the second round of amplification by PCR, **three nucleotide** sequences are amplified.

- This further reduces the pool of DNA fragments to a manageable level (< 100). **Autoradiography** can be performed for the **detection of DNA fragments.** Use of **radiolabeled primers and fluorescently labeled fragments** quickens AFLP.

Amplified fragment length polymorphism (AFLP)

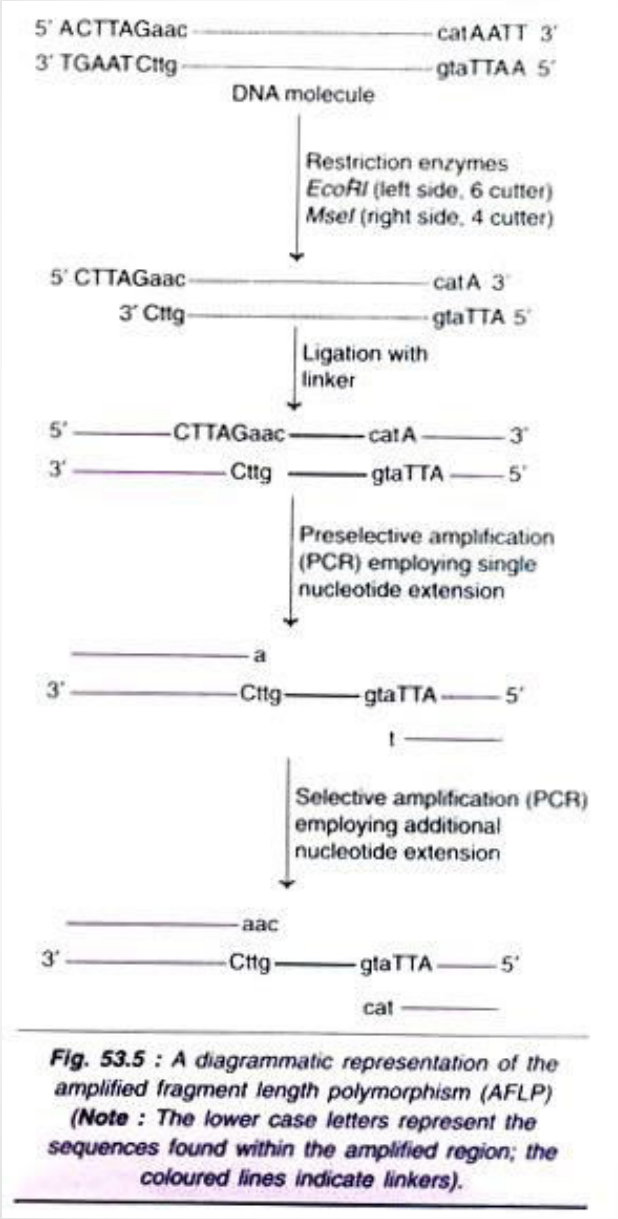
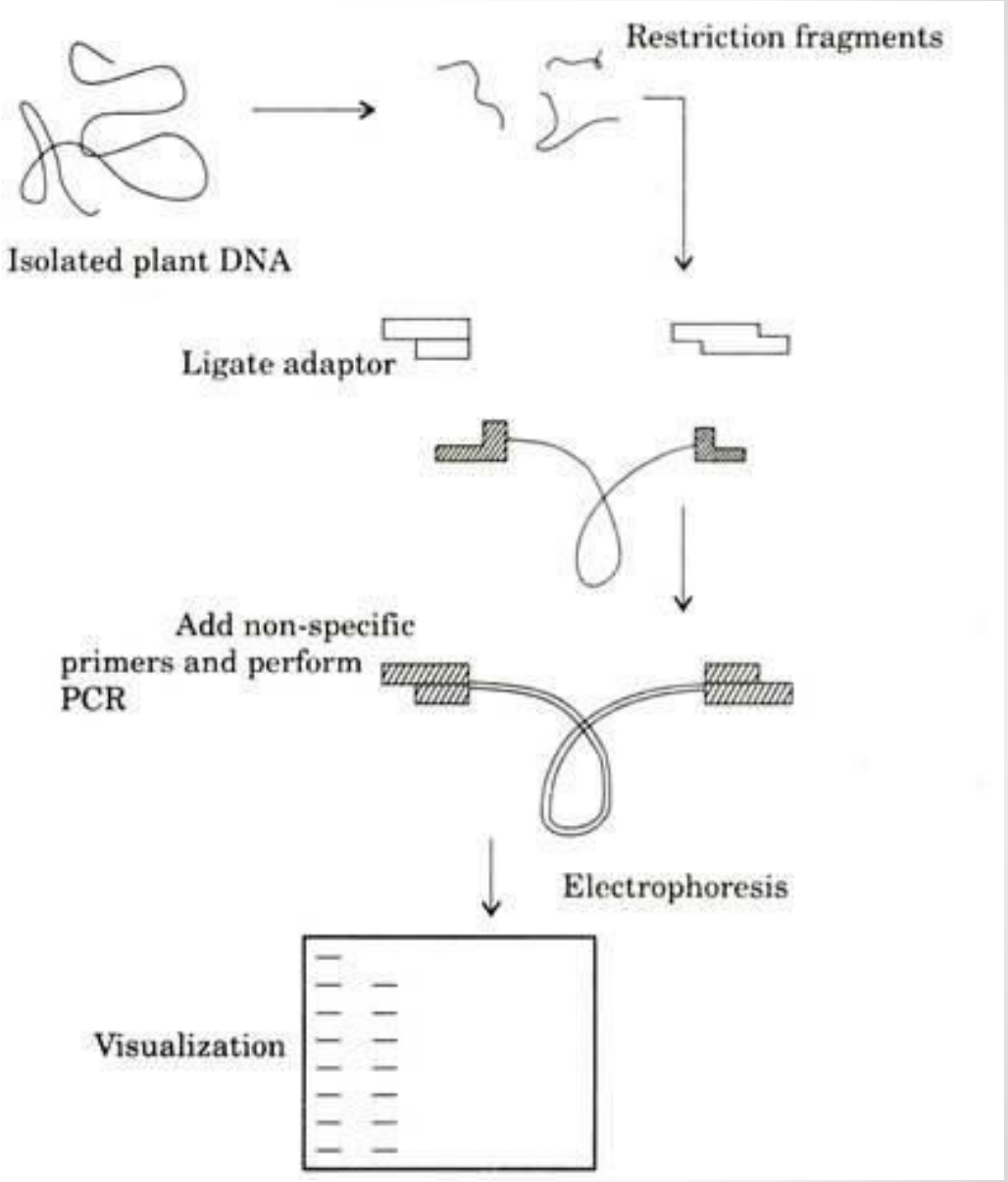


***Fig. 53.5** : A diagrammatic representation of the amplified fragment length polymorphism (AFLP) (**Note** : The lower case letters represent the sequences found within the amplified region; the coloured lines indicate linkers).*



Amplified fragment length polymorphism (AFLP)

- AFLP is primarily used in **genetic mapping**. Several economically important cereal crops such as **rice, barley and wheat** have been mapped by AFLP. The AFLP markers, which are produced by **different combinations** of **restriction enzymes**, are distributed throughout the genome. In **barley**, AFLP markers are located on the **long and short arm** of **all seven chromosomes.**

- These **AFLP markers** exhibit strong relation between the **number of markers** per chromosome and **length of the chromosome**. Similarly, in rice, cross between **Indica X Japonica** revealed that **50 AFLP markers** were located on every chromosome except in small chromosome 12. These polymorphic loci distributed throughout the genome of this species can be illustrated by AFLP technique.

Amplified fragment length polymorphism (AFLP)

- The existing barley map developed by **RFLP comprises of 157 RFLP loci** has been increased by **adding 118 AFLP markers**. The total map length was increased by 71% mainly attributed to gap filling, terminal extension and general expansion.

- The level of **polymorphism** detected in barley by **AFLP** can range from **12.2%** (between procter × Ny dinka) to **29.0%** (between L94 × voda). However, detection level of this polymorphic range is **comparatively less** than that of used by other mapping technique like RFLP.

- Increasing AFLP is used for several applications to assist the **rapid isolation** and **characterization** of **target genes.**

- **Breeding** for resistance is an important programme in **cereal research**. AFLP technique identified several markers closely linked to barley Mlo resistance gene. AFLP based fine mapping of the resistance gene locus to be delimited to 30 kb.

Amplified fragment length polymorphism (AFLP)

Merits:

i. **AFLP** is used to **screen large number** of **polymorphism**. It is possible to saturate specific regions of the genome for **map-based cloning** of target genes.

ii. AFLP has also been used in **phylogenetic studies** and distinguishing feature between **varieties**. In case of **barley studies**, varieties are grouped according to their **salt tolerance** and area of **origin by genotyping** with AFLP. Determination of ancestral origin of wheat variety using number of AFLP analysis which had previously not been possible using other molecular techniques due to the low genetic diversity of races.

iii. AFLP is also used to screen pools of **plasmid DNA** from **several clones**, enabling rapid isolation of genes tightly linked to markers.

iv. AFLP has recently been applied to the analysis of **quantitative traits in barley and rice.**

Amplified fragment length polymorphism (AFLP)

v. RFLP can be used to score semi-dominant markers. This was possible due to development of new software for **image analysis of fluorescent PCR products developed by key gene**. This was probably developed for use with AFLP and enable AFLP to score **semi-dominant marker.**

vi. AFLP is rapidly becomes preferred molecular technique for several different investigation particularly in **many areas of research.**

Amplified fragment length polymorphism (AFLP)

Demerits:

i. Not reliable to convert **AFLP into SCAR**

ii. Null allele cannot be detected

iii. Proprietary technique

iv. Require **high amount of DNA** than required in RAPD.

v. Relatively expensive technique owing to requirement of **silver staining** and **radio or non-radiolabelling.**

Markers Based on PCR Amplification

Sequence tagged sites (STS):

- Sequence tagged sites represent **unique simple copy segments of genomes**, whose **DNA sequences** are known, and which can be **amplified** by using **PCR**. STS markers are based on the **polymorphism** of **simple nucleotide repeats** e.g. $(GA)_n$, $(GT)_n$, $(CAA)_n$ etc. on the genome. STS have been recently developed in **plants**.
- When the **STS loci** contain **simple sequence length polymorphisms** (SSLPs), they are highly **valuable as molecular markers**. STS loci have been analysed and studied in a number of plant species.

Microsatellites:

Microsatellites are the **tandemly repeated** multi-copies of **mono-, di-, tri- and tetra nucleotide motifs**. In some instances, the **flanking sequence** of the **repeat sequences** may be unique. Primers can be designed for such flanking sequences to detect the **sequence tagged microsatellites** (STMS). This can be done by **PCR.**

Sequence characterized amplified regions (SCARs):

SCARs are the modified forms of **STS markers**. They are developed by **PCR primer** that are made for the ends of **RAPD fragment.** The **STS-converted RAPD markers** are sometimes referred to as **SCARs**. SCARs are useful for the **rapid development of STS markers.**

Markers Based on PCR Amplification

Sequence characterized amplified regions (SCARs):
SCARs are the modified forms of **STS markers**. They are developed by **PCR primer** that are made for the ends of **RAPD fragment.** The **STS-converted RAPD markers** are sometimes referred to as **SCARs**. SCARs are useful for the **rapid development of STS markers.**

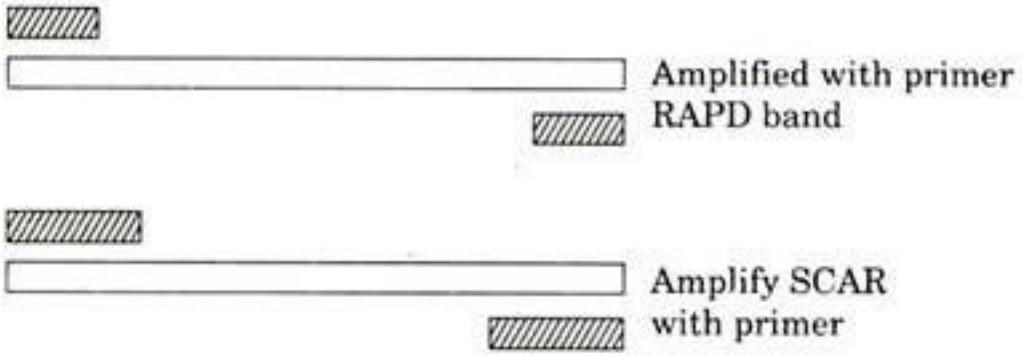


Fig. 24.3 Sequence characterised amplified region

Molecular Marker Assisted Selection

- Selection of the **desired traits and improvement of crops** has been a part of the **conventional breeding programmes**. This is predominantly based on the identification of **phenotypes**.
- It is now an accepted fact that the **phenotypes** do not necessarily represent the **genotypes**. Many a times the **environment** may mark the **genotype**.
- The **molecular marker assisted selection** is based on the **identification of DNA markers** that link/represent the **plant traits**. These traits include **resistance to pathogens and insects, tolerance to abiotic stresses, and various other qualitative and quantitative traits**.
- The advantage with a molecular marker is that a plant breeder can select a **suitable marker** for the **desired trait** which can be detected well in **advance**. Accordingly, breeding programmes can be planned.

The following are the major requirements for the molecular marked selection in plant breeding:

i. The marker should be **closely linked** with the **desired trait.**

ii. The marker screening methods must be **efficient, reproducible and easy to carry out.**

iii. The analysis should be **economical.**

Molecular Breeding

Molecular Breeding:

- With rapid progress in molecular biology and genetic engineering, there is now a possibility of improving the crop plants with respect to **yield and quality.** The term molecular breeding is frequently used to represent the **breeding methods** that are coupled with **genetic engineering techniques**.

- Improved agriculture to meet the food demands of the world is a high priority area. For several years, the conventional plant breeding programmes (**although time consuming**) have certainly helped to improve **grain yield and cereal production.**

- The development of **dwarf and semi-dwarf varieties** of **rice and wheat** have been responsible for the **'Green Revolution',** which has helped to feed millions of poverty-stricken people around the world. Many developments on the agriculture front are expected in the coming years as a result of **molecular breeding**.

Linkage analysis

Linkage analysis:

- Linkage analysis basically deals with studies to correlate the link between the **molecular marker** and a **desired trait.** This is an important aspect of molecular breeding programmes.

- Linkage analysis has to be carried out among the populations of **several generations** to establish the appropriate linkage.

- In the earlier years, **linkage analysis** was carried out by use of **isoenzymes** and the associated **polymorphisms**. Molecular markers are now being used.

Quantitative Trait Loci

Quantitative Trait Loci (QTL):

- These are **many characteristics** controlled by **several genes** in a **complex manner**. Some good examples are **growth habit, yield, adaptability to environment, and disease resistance.**
- These are referred to as quantitative traits. The **locations** on the chromosomes for these genes are regarded as quantitative trait loci (**QTL**).
- The major problem, the plant breeder faces is how to **improve the complex character** controlled by many genes. It is not an easy job to manipulate multiple genes in **genetic engineering**. Therefore, it is a very difficult and time consuming process.
- For instance, development of **Golden Rice** (with enriched **pro-vitamin A**) involving the insertion of just **three genes** took about **seven years.**

Thank you