



Marker-Assisted Selection: Accelerating the Breeding Process

Sanjay Singh¹
Subaran Singh²
Supriya Ambawat³
Yogendra Singh⁴

¹Deptt. of Vegetable Science,
College of Horticulture, Maharana
Pratap Horticultural University
Karnal- 132001 (Haryana)

²Deptt. of Basic Sciences, College
of Horticulture, Maharana Pratap
Horticultural University
Karnal- 132001 (Haryana)

³Directorate of Research,
Agriculture University
Jodhpur -342304 (Rajasthan)

⁴Assistant Professor (Senior Scale)
Biotechnology, Department of
Genetics and Plant Breeding,
JNKVV, Jabalpur-482004 (M.P)



Open Access

*Corresponding Author
Subaran Singh*

Article History

Received: 14.08.2026

Revised: 19.08.2026

Accepted: 24.08.2026

This article is published under the
terms of the [Creative Commons
Attribution License 4.0](https://creativecommons.org/licenses/by/4.0/).

INTRODUCTION

Marker-assisted selection (MAS) is the use of DNA markers, in place of the visible expression of a trait, to decide which plants are retained in a breeding programme. Its principal advantage is the reduction in breeding cycle time. By conventional methods, the transfer of a single disease resistance gene into an acceptable commercial variety requires twelve to fourteen years. The gene is usually available only in a wild relative or in a landrace of poor agronomic value, and its transfer into an elite background requires repeated crossing and backcrossing, with every cycle of selection depending on expression of the trait in the field. Selection for disease resistance therefore depends on the occurrence of adequate disease pressure, such as whitefly-transmitted leaf curl virus in tomato, anthracnose in chilli or Fusarium wilt in banana.

DNA markers remove this dependence. A marker is a short stretch of DNA that differs in sequence between the two parents and lies close enough to the gene of interest for the two to be inherited together, so that the gene can be traced in the laboratory instead of being awaited in the field. A small leaf disc is taken from a fifteen-day-old seedling, its DNA is extracted and amplified by PCR, and the genotype of that plant at the target gene is known within two days. Neither disease pressure, nor an epidemic, nor a complete crop season is required. Seedling screening of this kind is practised for the Ty-2 and Ty-3 genes against leaf curl virus in tomato and for Fom-2 against Fusarium wilt in melon.

Limitations of conventional breeding

Selection based on visual observation has two major limitations, both of which prolong the breeding cycle. The first is that the plant must reach the stage at which the trait is expressed.

Fruit firmness in tomato, keeping quality in onion, seed number in citrus, flower colour and vase life in chrysanthemum, every one of these needs a plant carried through to harvest or to full bloom. The second is that the trait must be induced before it can be recorded. A late blight screening block in potato yields no useful information in a season of low inoculum, since

plants that escape infection are indistinguishable from plants that resisted it. Recessive genes present a further difficulty, because a plant carrying one copy is phenotypically identical to a plant carrying none, and the two can be separated only by selfing and scoring the next generation.

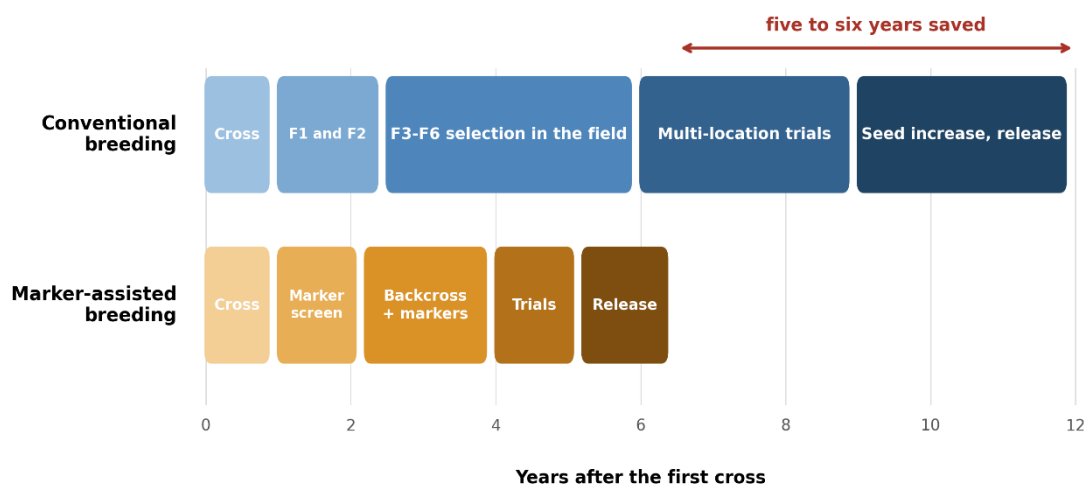


Fig. 1: Time taken by conventional breeding and marker-assisted breeding. Markers shorten the middle part of the programme and not the trial and release stages at the end.

Marker

A marker is any easily scored, heritable feature that is inherited together with a gene which cannot itself be observed directly. The marker and the gene lie close together on the same chromosome and are therefore transmitted to the progeny as a unit, so that the presence of the marker indicates the presence of the gene. The marker itself need not contribute to the trait; its value lies entirely in this linkage, and selection for the marker constitutes indirect selection for the gene. Three classes of marker have been used in plant breeding. Morphological markers were the earliest and include leaf shape, hypocotyl colour, dwarfness and flower colour. They require no cost to score, but only a small number are available in any crop, they are modified by weather and management, and most represent traits of no value in the variety. Biochemical markers followed, chiefly isozymes such as peroxidases and esterases separated on starch or polyacrylamide gels.

They are little influenced by the environment and remain in use for varietal identification and hybrid purity testing, but the thirty or so usable systems are too few to cover all the chromosomes. DNA or molecular markers constitute the third class. Each is a short stretch of DNA that differs in sequence between the two parents and is detected in the laboratory. Their number is effectively unlimited, they can be scored in a fifteen-day-old seedling, and the result is identical in root, leaf or petal tissue, irrespective of season.

Features of a useful marker

Not every difference in DNA is useful. A marker earns its place when it is polymorphic, meaning it genuinely differs between the two parents being crossed; when it is reproducible, so the same plant gives the same band in Hisar and in Bengaluru; when it is codominant, so a heterozygote can be told apart from both homozygotes; when it is cheap enough to run on several hundred plants; and above all when

it sits close to the gene. A marker that satisfies the first four conditions but fails the last is unsuitable for selection, since it yields confident but incorrect decisions.

Dominant and codominant markers

This classification determines how much information a single band provides. A dominant marker either produces a band or does not. A plant with two copies of the allele and a plant with one copy both show it, so the homozygote cannot be separated from the heterozygote without growing another generation. A codominant marker gives one band size from the resistant parent's allele and a different size from the susceptible one and a heterozygote shows both. In a backcross programme, where nearly every plant of interest is heterozygous, this distinction is decisive.

Closeness of the marker to the gene

Distance along a chromosome is measured in centiMorgans and one centiMorgan means roughly one crossover per hundred gametes. A marker one centiMorgan from its gene parts company with it in about one plant per hundred each generation, which is acceptable in most breeding programmes. A marker ten centiMorgans away gives an incorrect result in about one plant in ten, and after two or three rounds of selection the nursery accumulates plants that carry the band but not the gene. This is why breeders prefer flanking markers, one on either side of the target, so a plant is accepted only when both bands appear and losing the gene needs two crossovers rather than one. A further improvement is a marker developed from within the gene itself, termed a functional or gene-based marker, which cannot be separated from the gene by recombination because it forms part of the gene sequence.

Main types of DNA markers

DNA markers are classified according to the physical basis on which the polymorphism is detected. Broadly there are three groups:

hybridisation-based markers, PCR-based markers and the newer sequence-based ones. RFLP belongs to the first and came earliest, cutting DNA with restriction enzymes and blotting the fragments; it is accurate and codominant, but slow and requires a large quantity of high-quality DNA. Most marker systems developed subsequently are based on PCR.

Within the PCR group the distinction that matters are whether the primer is arbitrary or designed. RAPD and ISSR use arbitrary or semi-arbitrary primers and sample the genome at random, which makes them rapid and inexpensive for diversity studies but unsuitable for routine selection, because a faint band obtained on one occasion may not be reproduced on another. Markers based on designed primers are reproducible. SSRs exploit the fact that short motifs such as GA or CTT repeat a variable number of times at thousands of places in every genome and that the repeat number differs between genotypes. SCAR markers take an unreliable RAPD band, sequence its ends and design a long specific primer pair that amplifies only that band. CAPS markers amplify a fragment and then cut it with a restriction enzyme, so a single letter difference becomes a visible difference in banding and STS markers turn any mapped sequence into the same kind of repeatable test.

The sequence-based group represents the present direction of the field. SNP markers read a single letter of the sequence and because single-letter differences are everywhere, they are effectively unlimited in number. They are scored in fluorescent plate assays or on chips rather than on gels, which is what makes it possible to test a few thousand plants overnight. Genotyping-by-sequencing goes further and reads thousands of sites in one run, giving a breeder the target gene and a picture of the whole genetic background at the same time.

Table 1: Marker systems commonly used in a crop breeding laboratory.

Marker	What it detects	Reading	Practical note
RFLP	Change in a restriction enzyme cutting site, seen after blotting	Codominant	Accurate but slow; requires much high-quality DNA
RAPD	Random short primers amplifying wherever they fit	Dominant	Rapid and inexpensive, but poorly reproducible between laboratories
AFLP	Restriction digestion followed by selective amplification	Mostly dominant	Very many bands at once, good for diversity work
SSR	How many times a short motif such as GA or CTT repeats	Codominant	Most widely used in Indian breeding laboratories
SCAR / CAPS	A specific sequence turned into a single reliable band	Either	Built from a mapped marker, made for routine selection
SNP	One letter of the sequence differing between parents	Codominant	Read in plates or chips; thousands of plants per run

Steps followed in the laboratory

1. Taking out the DNA from the leaf

A hundred milligrams of young leaf is ground in a buffer, usually the CTAB method or a kit version of it. The detergent breaks the cell and nuclear membranes, salt keeps proteins in solution, chloroform pulls the proteins away and cold isopropanol brings the DNA out as a white thread. One seedling gives enough DNA for dozens of marker tests and the leaf that was taken regrows.

2. Copying the required piece of DNA

The polymerase chain reaction amplifies only the region of the genome lying between two

primers. Two primers, each about twenty bases long, are designed to match the sequence on either side of the marker region. In the thermal cycler the mixture is heated to about 94 degrees so the double helix comes apart, cooled to between 50 and 60 degrees so the primers can sit down on their matching sequences, then warmed to 72 degrees so that Taq polymerase, an enzyme from a hot spring bacterium that survives the boiling step, builds a new strand from each primer. Every cycle doubles the copies and thirty-five cycles turn one molecule into roughly a billion.

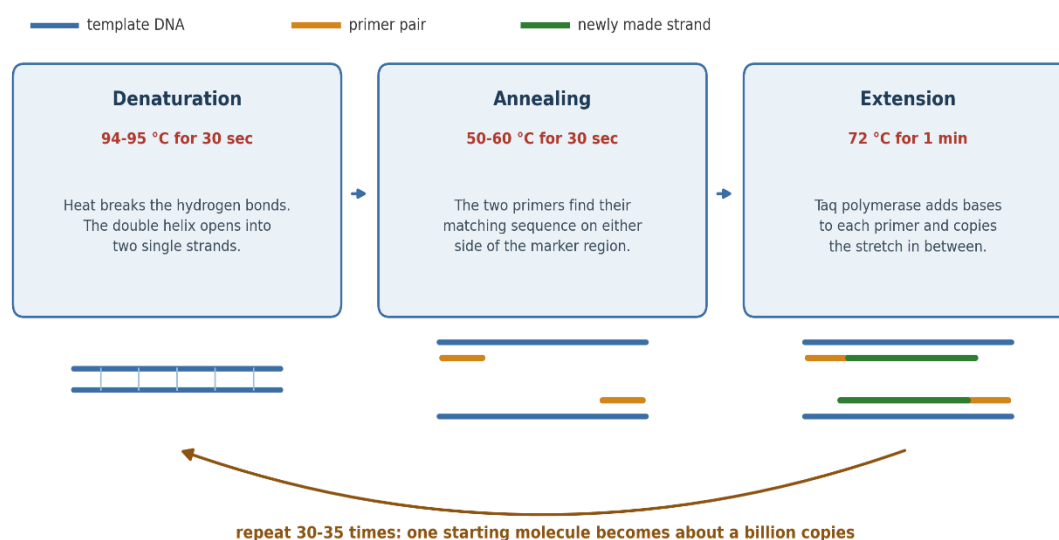


Fig. 2: The three temperature steps of one PCR cycle. Only the part lying between the two primers is copied and this makes the test specific.

The amplified product is abundant but cannot be seen directly. Agarose gel electrophoresis both renders it visible and measures its size. Molten agarose is poured into a tray with a comb and when it sets the comb leaves a row of small wells. The gel goes into a tank of buffer with an electrode at each end, samples are loaded and the current is switched on. DNA carries a negative charge because of its

phosphate backbone, so every fragment sets off towards the positive electrode and the mesh of the gel sorts them: a short fragment slips through and travels far, a long one is held back near the well. After thirty to sixty minutes a dye bound to the DNA glows under ultraviolet light and a lane of fragments of known size, the ladder, lets each band be given a number in base pair

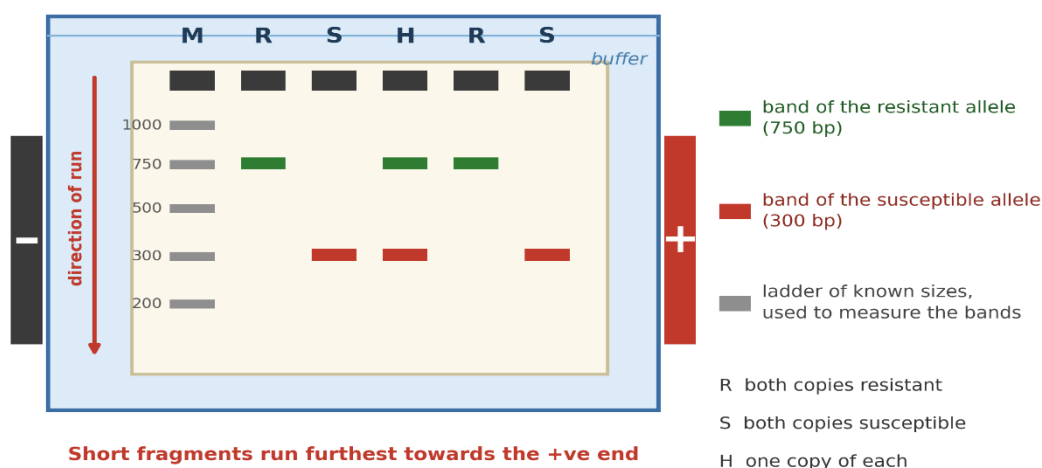


Fig. 3: Separation of DNA fragments on an agarose gel. The gel sorts the fragments by size and the position of a band shows which allele the plant carries.

4. Reading the bands and selecting the plants

The banding pattern on the gel is then read genotype by genotype. A plant showing only the upper band carries two copies of one allele,

a plant showing only the lower band carries two copies of the other and a plant showing both is a heterozygote with one of each. The breeder now knows which seedlings to keep.



Fig. 4: The complete process from a seedling tray to a selection decision in about two days.

Marker-assisted selection does not replace field evaluation. The marker establishes what the plant carries, while multi-location trials establish whether that combination is worth releasing. What the marker changes is the stage at which the selection decision can be taken. Discarding nine hundred out of a thousand

seedlings in the second week of the season, on the basis of genotype rather than on expectation, recovers the seasons formerly spent in identifying the carriers, and this is the source of the reduction in breeding cycle length shown in Fig. 1.