

Marker assisted selection and Crop improvement

Marker assisted selection or marker aided selection (MAS) is a process whereby a marker (morphological, biochemical or one based on DNA/RNA variation) is used for indirect selection of a genetic determinant or determinants of a trait of interest (i.e. productivity, disease resistance, abiotic stress tolerance, and/or quality). This process is used in plant and animal breeding.

Considerable developments in biotechnology have led plant breeders to develop more efficient selection systems to replace traditional phenotypic-pedigree-based selection systems.

Marker assisted selection (MAS) is **indirect selection process** where a trait of interest is selected not based on the trait itself but on a marker linked to it. For example if MAS is being used to select individuals with a disease, the level of disease is not quantified but rather a marker allele which is linked with disease is used to determine disease presence. The assumption is that linked allele associates with the gene and/or quantitative trait locus (QTL) of interest. MAS can be useful for traits that are difficult to measure, exhibit low heritability, and/or are expressed late in development.

Marker types

A marker may be:

- **Morphological** - First markers loci available that have obvious impact on morphology of plant. Genes that affect form, colouration, male sterility or resistance among others have been analyzed in many plant species. Examples of this type of marker may include the presence or absence of awn, leaf sheath colouration, height, grain colour, aroma of rice etc. In well-characterized crops like maize, tomato, pea, barley or wheat, tens or even hundreds of such genes have been assigned to different chromosomes.
- **Biochemical**- A gene that encodes a protein that can be extracted and observed; for example, isozymes and storage proteins.
- **Cytological** - The chromosomal banding produced by different stains; for example, G banding.

- **Biological-** Different pathogen races or insect biotypes based on host pathogen or host parasite interaction can be used as a marker since the genetic constitution of an organism can affect its susceptibility to pathogens or parasites.
- **DNA-based and/or molecular-** A unique (DNA sequence), occurring in proximity to the gene or locus of interest, can be identified by a range of molecular techniques such as RFLPs, RAPDs, AFLP, DAF, SCARs, microsatellites etc.

Sax in 1923 first reported association of a simply inherited genetic marker with a quantitative trait in plants when he observed segregation of seed size associated with segregation for a seed coat colour marker in beans (*Phaseolus vulgaris* L.). Rasmusson in 1935 demonstrated linkage of flowering time (a quantitative trait) in peas with a simply inherited gene for flower colour.

Gene vs marker

The gene of interest is directly related with production of protein(s) that produce certain phenotypes whereas markers should not influence the trait of interest but are genetically linked (and so go together during segregation of gametes due to the concomitant reduction in homologous recombination between the marker and gene of interest). In many traits genes are discovered and can be directly assayed for their presence with a high level of confidence. However, if a gene is not isolated marker's help is taken to tag a gene of interest. In such case there may be some false positive results due to recombination between marker of interest and gene (or QTL). A perfect marker would elicit no false positive results.

Important properties of ideal markers for MAS

An ideal marker:

- Easy recognition of all possible phenotypes (homo and heterozygotes) from all different alleles
- Demonstrates measurable differences in expression between trait types and/or gene of interest alleles, early in the development of the organism
- Has no effect on the trait of interest that varies depending on the allele at the marker loci
- Low or null interaction among the markers allowing the use of many at the same time in a segregating population

- Abundant in number
- Polymorphic

Demerits of morphological markers

Morphological markers are associated with several general deficits that reduce their usefulness including:

- the delay of marker expression until late into the development of the organism
- dominance
- deleterious effects
- pleiotropy
- confounding effects of genes unrelated to the gene or trait of interest but which also affect the morphological marker (epistasis)
- rare polymorphism
- frequent confounding effects of environmental factors which affect the morphological characteristics of the organism

To avoid problems specific to morphological markers, the DNA-based markers have been developed. They are highly polymorphic, simple inheritance (often codominant), abundantly occur throughout the genome, easy and fast to detect, minimum pleiotropic effect and detection is not dependent on the developmental stage of the organism. Numerous markers have been mapped to different chromosomes in several crops including rice, wheat, maize, soybean and several others. Those markers have been used in diversity analysis, parentage detection, DNA fingerprinting, and prediction of hybrid performance. Molecular markers are useful in indirect selection processes, enabling manual selection of individuals for further propagation.

Selection for major genes linked to markers

The major genes which are responsible for economically important characteristics are frequent in the Plant Kingdom. Such characteristics include disease resistance, male sterility, self-incompatibility; others related to shape, colour, and architecture of whole plants and are often of mono- or oligogenic in nature. The marker loci which are tightly linked to major genes can be used for selection and are sometimes more efficient than direct selection for the target gene. Such advantages in efficiency may be due for example, to higher expression of the marker mRNA in such cases that the marker is actually a gene. Alternatively, in such cases that the target gene of interest differs between two alleles by a

difficult-to-detect single nucleotide polymorphism, an external marker (be it another gene or a polymorphism that is easier to detect, such as a short tandem repeat) may present as the most realistic option.

Situations that are favorable for molecular marker selection

There are several indications for the use of molecular markers in the selection of a genetic trait.

In such situations that:

- the selected character is expressed late in plant development, like fruit and flower features or adult characters with a juvenile period (so that it is not necessary to wait for the organism to become fully developed before arrangements can be made for propagation)
- the expression of the target gene is recessive (so that individuals which are heterozygous positive for the recessive allele can be crossed to produce some homozygous offspring with the desired trait)
- there is requirement for the presence of special conditions in order to invoke expression of the target gene(s), as in the case of breeding for disease and pest resistance (where inoculation with the disease or subjection to pests would otherwise be required). This advantage derives from the errors due to unreliable inoculation methods and the fact that field inoculation with the pathogen is not allowed in many areas for safety reasons. Moreover, problems in the recognition of the environmentally unstable genes can be eluded.
- the phenotype is affected by two or more unlinked genes (epistasis). For example, selection for multiple genes which provide resistance against diseases or insect pests for gene pyramiding.

The cost of genotyping (an example of a molecular marker assay) is reducing while the cost of phenotyping is increasing particularly in developed countries thus increasing the attractiveness of MAS as the development of the technology continues.

Steps for MAS

Generally the first step is to map the gene or quantitative trait locus (QTL) of interest first by using different techniques and then use this information for marker assisted selection. Generally, the markers to be used should be close to gene of interest (<5 recombination unit or cM) in order to ensure that only minor fraction of the selected individuals will be

recombinants. Generally, not only a single marker but rather two markers are used in order to reduce the chances of an error due to homologous recombination. For example, if two flanking markers are used at same time with an interval between them of approximately 20cM, there is higher probability (99%) for recovery of the target gene.

QTL mapping techniques

In plants QTL mapping is generally achieved using bi-parental cross populations; a cross between two parents which have a contrasting phenotype for the trait of interest are developed. Commonly used populations are recombinant inbred lines (RILs), doubled haploids (DH), back cross and F₂. Linkage between the phenotype and markers which have already been mapped is tested in these populations in order to determine the position of the QTL. Such techniques are based on linkage and are therefore referred to as "linkage mapping".

Single step MAS and QTL mapping

In contrast to two-step QTL mapping and MAS, a single-step method for breeding typical plant populations has been developed.^[5] In such an approach, in the first few breeding cycles, markers linked to the trait of interest are identified by QTL mapping and later the same information is used in the same population. In this approach, pedigree structures are created from families that are created by crossing number of parents (in three-way or four way crosses). Both phenotyping and genotyping is done using molecular markers mapped the possible location of QTL of interest. This will identify markers and their favorable alleles. Once these favorable marker alleles are identified, the frequency of such alleles will be increased and response to marker assisted selection is estimated. Marker allele(s) with desirable effect will be further used in next selection cycle or other experiments.

High-throughput genotyping techniques

Recently high-throughput genotyping techniques are developed which allows marker aided screening of many genotypes. This will help breeders in shifting traditional breeding to marker aided selection. One of example of such automation is using DNA isolation robots, capillary electrophoresis and pipetting robots.

One of recent example of capillary system is Applied Biosystems 3130 Genetic Analyzer. This is the latest generation of 4-capillary electrophoresis instruments for the low to medium throughput laboratories.

Use of MAS for backcross breeding

A minimum of five or six-backcross generations are required to transfer a gene of interest from a donor (may not be adapted) to a recipient (recurrent – adapted cultivar). The recovery of the recurrent genotype can be accelerated with the use of molecular markers. If the F1 is heterozygous for the marker locus, individuals with the recurrent parent allele(s) at the marker locus in first or subsequent backcross generations will also carry a chromosome tagged by the marker.

Marker assisted gene pyramiding

Gene pyramiding has been proposed and applied to enhance resistance to disease and insects by selecting for two or more than two genes at a time. For example in rice such pyramids have been developed against bacterial blight and blast. The advantage of use of markers in this case allows selecting for QTL-allele-linked markers that have same phenotypic effect.