Introduction

The development of molecular marker technology has caused renewed interest in genetic mapping. An appropriate mapping population, suitable marker system and the software for analyses of data are the key requirements for a molecular mapping and molecular breeding programme. Genetic map construction requires that the researchers: (i) select the most appropriate mapping population(s); (ii) calculate pairwise recombination frequencies using these population; (iii) establish linkage groups and estimate map distances; and (iv) determine map order.

Since large mapping populations are often characterized by different marker systems, map construction has been computerized. Computer software packages, such as Linkage1 (Suiter et al., 1983), GMendel (Echt et al., 1992), Mapmaker (Lander and Botstein, 1986; Lander et al., 1987), Mapmanager (Manly and Eliot, 1991) and Joinmap (Stam, 1993), have been developed to aid in the analysis of genetic data for map construction. These programmes use data obtained from the segregating populations to estimate recombination frequency that are then used to determine the linear arrangement of genetic markers.

Mapping Populations

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. When the objective is to search for genes controlling a particular trait, genetic variation of trait between parents is important. If the parents are greatly different at phenotypic level for a trait, there is a reasonable chance that genetic variation exists between the parents, although uncontrolled environmental effects could create large phenotypic variation without any genetic basis for the effects. However, lack of phenotypic variation between parents does not mean that there is no genetic variation, as different sets of genes could result in same phenotype.

Selection of parents for developing mapping population: 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. Consideration must be given to the source of parents (adapted vs
exotic) used in developing mapping population. Chromosome pairing and recombination rates can be severely disturbed (suppressed) in wide crosses and generally yield greatly reduced linkage distances (Albine and Jones, 1987; Zamir and Tadmar, 1986). Wide crosses will usually provide segregating populations with a relatively large array of polymorphism when compared to progeny segregating in a narrow cross (adapted x adapted). To have significant value in crop improvement programme, a map made from a wide cross must be collinear (i.e. order of loci similar) with map constructed using adapted parents.

**Types of mapping populations:** Different types of mapping populations that are often used in linkage mapping are: (i) F$_2$ population; (ii) F$_2$ derived F$_3$ (F$_2$:F$_3$) populations; (iii) Backcrosses; (iv) Doubled haploids (DHs); (v) Recombinant Inbred Lines (RILs); and (vi) Near-isogenic Lines (NILs). The characteristic features, merits and demerits of each of these populations are briefly presented below:

**F$_2$ population:**
- Produced by selfing or sib mating the individuals in segregating populations generated by crossing the selected parents.
- F$_2$ individuals are products of single meiotic cycle
- Ratio expected for dominant marker is 3:1 and for codominant marker is 1:2:1

**Merits**
- Best population for preliminary mapping
- Requires less time for development
- Can be developed with minimum efforts, when compared to other populations

**Demerits**
- Linkage established using F$_2$ population is based on one cycle of meiosis
- F$_2$ populations are of limited use for fine mapping
- Quantitative traits cannot be precisely mapped using F$_2$ population as each individual is genetically different and cannot be evaluated in replicated trials over locations and years. Thus, the effect the G x E interaction on the expression of quantitative traits cannot be precisely estimated.
- Not a long-term population; impossible to construct exact replica or increase seed amount

**F$_2$ derived F$_3$ (F$_2$:F$_3$) population:**
- F$_2$:F$_3$ population is obtained by selfing the F$_2$ individuals for a single generation
- Suitable for specific situations like
  - Mapping quantitative traits
  - Mapping recessive genes
- The F$_2$:F$_3$ family can be used for reconstituting the genotype of respective F$_2$ plants, if needed, by pooling the DNA from plants in the family

**Demerit:** Like F$_2$ population, it is not ‘immortal’

**Backcross Mapping Population:**
Backcross populations are generated by crossing the F$_1$ with either of the parents. Usually in genetic analysis, backcross with recessive parent (testcross) is used. With respect to molecular markers, the backcross with dominant parent (B$_1$) would segregate in a ratio 1:0 and 1:1 for dominant and codominant markers, respectively. However, backcross with recessive parent (B$_2$) or testcross would segregate in a ratio of 1:1 irrespective of the nature of marker. Like an F$_2$ population, the backcross populations require less time to be developed, but are not ‘immortal’. However, the recombination information in case of backcrosses is based on only one parent (the F$_1$). The specific advantage of backcross populations is that, the populations can be further utilized for marker-assisted backcross breeding.

**Doubled Haploids (DHs):**
- Chromosome doubling of anther culture derived haploid plants from F$_1$ generates DHs. The suitability of doubled–haploid progenies for mapping project has been demonstrated in by Lefebvre et al. (1995) in pepper.
- DHs are also products of one meiotic cycle, and hence comparable to F$_2$ in terms of recombination information.
- The expected ratio for the marker is 1:1, irrespective of genetic nature of marker (whether dominant or codominant).

**Merits**
- DHs are permanent mapping population and hence can be replicated and evaluated over locations and years and maintained without any genotypic change
- Useful for mapping both qualitative and quantitative characters
- Instant production of homozygous lines, thus saving time

**Demerits**
- Recombination from the male side alone is accounted.
- Since it involves *in vitro* techniques, relatively more technical skills are required in comparison with the development of other mapping populations
- Often suitable culturing methods / haploid production methods are not available for number crops and different crops differ significantly for their tissue culture response. Further, anther culture induced variability should be taken care of.

**Recombinant Inbred Lines (RILs):**
- RILs are produced by continuous selfing or sib mating the progeny of individual members of an F$_2$ population until complete homozygous is achieved
- SSD method is best suited for developing RILs. Bulk method and pedigree methods without selection can also be used
- RILs also equalize marker types like DHs, the genetic segregation ratio for both dominant and co dominant marker would be 1:1
- RILs developed though brother-sister mating require more time than those developed through selfing. The number of inbred lines required is twice, in case they are developed through brother–sister mating compared to selfing particularly, when linkage is not very tight.
**Merits**
- Once homozygosity is achieved, RILs can be propagated indefinitely without further segregation
- Since RILs are immortal population, they can be replicated over locations and years and therefore are of immense value in mapping QTLs
- RILs being obtained after several cycles of meiosis, are very useful in identifying tightly linked makers
- RIL populations obtained by selfing have twice the amount of observed recombination between very closely linked markers as compared to population derived from a single cycle of meiosis

**Demerits**
- Requires many seasons / generations to develop
- Developing RILs is relatively difficult in crops with high inbreeding depression

**Immortalized F₂ Population:**

Immortalized F₂ populations can be developed by paired crossing of the randomly chosen RILs derived from a cross in all possible combinations excluding reciprocals. The set of RILs used for crossing along with the F₁s produced, provide a true representation of all possible genotype combinations (including the heterozygotes) expected in the F₂ of the cross from which the RILs are derived. The RILs can be maintained by selfing and required quantity of F₁ seed can be produced at will by fresh hybridization. This population therefore provides an opportunity to map heterotic QTLs and interaction effects from multilocation data.

**Near-Isogenic Lines (NILs):**
- NILs are generated either by repeated selfing or backcrossing the F₁ plants to the recurrent parents.
- NILs developed through backcrossing are similar to recurrent parent but for the gene of interest, while NILs developed though selfing are similar in pair but for the gene of interest (however, differ a lot with respect to the recurrent parent)
- Expected segregation ratio of the markers is 1:1 irrespective of the nature of marker

**Merits**
- Like DHs and RILs, NILs are also ‘immortal mapping population’
- Suitable population for tagging the trait, wherever such population is available
- NILs are quite useful in functional genomics

**Demerits**
- Require many generations for development
- Directly useful only for molecular tagging of the gene concerned, but not for linkage mapping
- Linkage drag is a potential problem in constructing NILs, which has to be taken care of.

**Bulk Segregant Analysis**
Besides the above-mentioned populations, Bulk Segregant Analysis (BSA) approach, using any one of the above-mentioned populations (except NILs) is frequently used in gene tagging. BSA is based on the principle of isogenic lines. In BSA, two parents (say a resistant and susceptible), showing high degree of molecular polymorphism and contrast for the target trait are crossed and F_1 is selfed to generate F_2 population. In F_2, individual plants are phenotyped for resistance and susceptibility. Usually, the DNA isolated from 10 plants in each group is pooled to constitute resistant and susceptible bulks. The resistant parent, susceptible parent, resistant bulk and susceptible bulk, are surveyed for polymorphism using polymorphic markers. A marker showing polymorphism between parents as well as bulks is considered putatively linked to the target trait, and is further used for mapping using individual F_2 plants. Conceptually, the genetic constitution of the two bulks is similar, but for the genomic region associated with the target trait. Hence, they serve the purpose of isogenic lines in principle.

It has been observed over experiments that when 10 plants are sampled in each group for constituting the bulk, the probability of a polymorphic marker (between parents as well as bulks) not being linked to the target trait is extremely low (10^{-19}). Hence, usually 10 plants are used for constituting the bulks. However, this number may vary depending upon the types of mapping populations used. In absence of isogenic lines, the BSA approach provides a very useful alternative for gene tagging (Michelmore et al., 1991).

**Combining Markers and Populations**

The genetic segregation ratio at maker locus is jointly determined by the nature of marker (dominant / codominant) and types of mapping populations (Table 1). Therefore, a thorough understanding of the nature of markers and mapping population is crucial for any mapping projects. Markers such RFLPs, microsatellites and CAPS etc. are codominant in nature, while AFLP, RAPD, ISSR are often scored as dominant markers. Mapping populations such as RILs and DHs equalize marker type because of fixation of parental alleles at marker locus in homozygous condition. These population result in 1: 1 segregation ratio at marker locus irrespective of genetic nature of markers, while an F_2 population segregates in 1: 2: 1 ratio for a codominant marker and in 3:1 ratio for dominant marker. Depending upon the segregation pattern, statistical analysis of marker data will vary.

*Table 1: Genetic segregation ratio at marker locus in different marker–population combinations.*

<table>
<thead>
<tr>
<th>Marker</th>
<th>Nature</th>
<th>Genetic Segregation Ratio</th>
<th>F_2</th>
<th>RILs</th>
<th>DHs</th>
<th>NILs</th>
<th>Backcross Popn.</th>
</tr>
</thead>
<tbody>
<tr>
<td>RFLP</td>
<td>Co-dominant</td>
<td></td>
<td>1:2:1</td>
<td>1:1</td>
<td>1:1</td>
<td>1:1</td>
<td>1:1 1:1</td>
</tr>
<tr>
<td>RAPD</td>
<td>Dominant</td>
<td>3:1</td>
<td>1:1</td>
<td>1:1</td>
<td>1:1</td>
<td>1:1</td>
<td>1:0 1:1</td>
</tr>
<tr>
<td>AFLP</td>
<td>Dominant</td>
<td>3:1</td>
<td>1:1</td>
<td>1:1</td>
<td>1:1</td>
<td>1:1</td>
<td>1:0 1:1</td>
</tr>
<tr>
<td>Microsatellites</td>
<td>Co-dominant</td>
<td>1:2:1</td>
<td>1:1</td>
<td>1:1</td>
<td>1:1</td>
<td>1:1</td>
<td>1:1 1:1</td>
</tr>
</tbody>
</table>

**Characterization of Mapping Populations**
Precise molecular and phenotypic characterization of mapping population is vital for success of any mapping project. Since the molecular genotype of any individual is independent of environment, it is not influenced by G x E interaction. However, trait phenotype could be influenced by the environment, particularly in case of quantitative characters. Therefore, it becomes important to precisely estimate the trait value by evaluating the genotypes in multilocation testing over years using immortal mapping populations to have a valid marker-trait association.

**Segregation Distortion of Markers in Linkage Mapping**

Significant deviation from expected segregation ratio in a given marker-population combination is referred to as segregation distortion. There are several reasons for segregation distortion, including: gamete/zygote lethality, meiotic drive/preferential segregation, sampling/selection during population development and differential responses of parental lines to tissue culture in case of DHs. Segregation distortion can also be specific with respect to some markers in an otherwise normal mapping population. It is therefore important that the ‘goodness of fit’ of segregation ratio must be tested for individual marker locus and if necessary, the marker showing high degree of segregation distortion be eliminated from the analysis.

**Choice of Mapping Populations**

It is evident from the foregoing discussion that the short-term mapping populations such as F_2_, backcross and conceptual near isogenic lines developed through BSA approach can be a good starting point in molecular mapping, while long-term mapping populations such as RILs, NILs and DHs must be developed and characterized properly with respect to the traits of importance for global mapping projects. As a matter of fact, the development and phenotypic characterization of mapping populations should become an integral part of the ongoing breeding programmes in important crops. At this point, the role of geneticists and plant breeders becomes crucial to reap the benefits of molecular plant breeding.

**Suggested Readings**


Development, Characterization and Utilization of Mapping Populations

Parent 1 → F₁ → F₂

Backcrossing

Backcross Population → Pedigree Method

Pedigree Method → NILs

NILs → Functional genomics

NILs → Characterization for target traits

Characterization for target traits → Identification of markers linked to target traits → MAŠ

F₂ → SSD

SSD → RILs

RILs → Molecular genotyping

F₂ → Bulk Method

Anther Culture

Backcross Population → DSM

DSM → Immortal mapping populations

Immortal mapping populations → Map based cloning of genes

Map based cloning of genes

Parent 2