JUST AGRICULTURE
multidisciplinary e-Newsletter

MAPPING POPULATIONS IN CROP IMPROVEMENT

B.L. Meena[*1], H.S. Meena[1], P.B. Singh[2] and S.P. Das[3]

[1]ICAR-DRMR, Bharatpur Rajasthan-321303, India, [2]MPUAT, Rajasthan College of Agriculture, Udaipur Rajasthan -313001, India,

[3]ICAR- RC for NEH Region Tripura Centre, Lembucherra

ARTICLE ID: 013

Abstract

A population that is suitable for linkage mapping of genetic markers is known as mapping population. Mapping populations are generated by crossing two or more genetically diverse lines and handling the progeny in a definite fashion. Mapping populations are used for determining genetic distances between pairs of loci/genes and to map them to specific locations in the genome. There are basically two types of mapping populations, viz., primary and secondary mapping populations. Primary mapping populations are created by hybridization between two homozygous lines usually having contrasting forms for the traits of interest. Secondary mapping populations are developed by crossing two lines/individuals selected from a mapping population; they are developed mainly for fine mapping of the genomic region of interest.

The primary mapping populations are of the following different types: (1) F_2, (2) F_2derived F_3 (F_2:F_3), (3) backcross (BC), (4) backcross inbred lines (BILs), (5) doubled haploids (DHs), (6) recombinant inbred lines (RILs), (7) near-isogenic lines (NILs),(8) chromosomal segment substitution lines (CSSLs), (9) immortalized F_2, (10) advanced intercross lines, (11) recurrent selection back(RSB) populations, and (12) interconnected populations

F_2 population: A F_2 mapping population comprises the progeny produced by selfing or sib-mating of the F_1 individuals from a cross between the selected parents. The F_1 individuals would be heterozygous for all the loci for which their parents differ from each other. Each F_2 individual is expected to have a unique combination of linkage blocks from the two parents, and this difference is the basis for detection of linkage between pairs of loci. Since F_2generation is the product of a single meiotic cycle (in the F_1plants), only one round of

JUST AGRICULTURE
multidisciplinary e-Newsletter

recombination can occur between any of two loci. Therefore, the estimates of recombination frequencies between pairs of loci obtained from F_2 populations serve as a reference point. In a F_2 population the ratios expected for dominant and codominant markers are 3:1 and 1:2:1, respectively. F_2 populations are the best suited for preliminary mapping of markers and oligogenes. The F_2 populations provide estimates of additive, dominance, and epistatic components of the genetic variance. These populations capture the recombination events from both male and female parents of the F_2 plants.

F_2-Derived F_3 Population: A F_2-derived F_3 or $F_{2:3}$population is obtained by selfing the F_2 individuals for a single generation and harvesting the seeds from each F_2plant separately so that each F_2 plant is represented as an individual plant progeny. The DNA for genotyping is obtained from individual F_2plants or it can be reconstructed from a bulk of at least 20 plants from each F_3 family since this bulked DNA may be expected to represent the genotype of the parental F_2 plant. F2:3 populations are suitable for mapping of oligogenic traits controlled by recessive genes and of QTLs since data can be recorded on multiple plants in each $F_{2:3}$family to compensate for sampling error.

Backcross Population: Backcross populations are generated by crossing F_1 plants with either of the two parents of the concerned F_1. Genetic analysis can be performed only when there is detectable phenotypic segregation for the target trait in the backcross generation. Therefore, the F_1 is, as a rule, backcrossed to the recessive parent, i.e., the parent having the recessive form of the target trait. Such a backcross is called testcross, is usually denoted by B_2, and exhibits 1:1 ratio for the trait phenotype, dominant molecular markers present in coupling phase with respect to the target trait, and codominant markers in either phase. However, it would show 1:0 ratio, i.e., no segregation, for dominant markers present in repulsion phase in relation to the target trait. In contrast, progeny from backcross with the dominant parent would display ratio for the trait phenotype and dominant markers present in coupling phase with respect to the target trait.

Doubled Haploids: Doubled haploid (DH) plants are obtained by chromosome doubling of haploid plants usually derived by culture of anthers/pollen grains produced by F_1 plants. Generally, colchicine is used to double the chromosome number of haploids, seeds from individual DH plants are harvested separately and maintained as DHlines. The expected ratio for the genes as well as markers in a DH population is 1:1 irrespective of the marker being

JUST AGRICULTURE
multidisciplinary e-Newsletter

dominant or codominant. DH populations, can be evaluated in replicated trials and are suitable for mapping both qualitative and quantitative characters. Only additive and additive x additive interaction genetic variances can be estimated from DH populations as they consist of only homozygous plants. Therefore, DH populations are not suitable for mapping heterosis QTLs.

Recombinant Inbred Lines: Recombinant inbred lines (RILs) are a set of homozygous lines produced by continuous inbreeding/selfing of individual F_2 plants. The SSD method is the best suited for developing RILs, but bulk procedure and pedigree method without selection can also be used. It is important that the generation advance is carried out under an optimal environment that affords equal survival of the various genotypes and does not impose a selection pressure against some genotypes. The SSD procedure is followed for five or more (usually >8) generations, during which one seed is harvested from each plant of the F_2 and the later generations and seeds from all the plants are composited and planted to raise the next generation. At the end of SSD procedure, seeds from each plant are harvested separately to obtain as many RILs as there are individual plants in the SSD population. The expected ratio of the two homozygotes in the population is l: l. RILs have been widely used for the development of molecular marker linkage maps; detection of markers linked with genes governing qualitative traits like race specific vertical disease resistance, seed or flower color, seed/fruit shape etc.; identification of markers associated with QTLs involved in the control of traits like horizontal disease resistance, yield, days to flowering/maturity etc.; mapping of genes and QTLs; and the integration of the gene/QTL maps with molecular marker maps.

Immortalized F_2 Population: The population of single cross F_1s produced by inter crossing a set of RILs in pairs or as per some other scheme is known as immortalized F_2 population. IF_2 populations can also be developed by paired crossing of the randomly chosen RILs derived from a cross in all possible combinations, excluding the reciprocals; in this approach, the single crosses together with the parental RILs would constitute the IF_2 population. An IF_2 population provides a true representation of all possible genotypes, including the heterozygotes, expected in the IF_2 of the cross from which the RILs were derived. IF_2 populations support replicated evaluation of F_2 genotypes over locations and permit detection and mapping of QTLs, including heterosis QTLs, and estimation of various epistatic effects.

JUST AGRICULTURE
multidisciplinary e-Newsletter

Near-Isogenic Lines: Near-isogenic lines (NILs) are pairs of homozygous lines that are identical in genotype, except for a single gene/locus. NILs are generally produced by backcross procedure. In which a donor parent (DP, a homozygous line having the trait/allele of interest) is crossed with a recurrent parent (RP, a homozygous line lacking this trait/allele), and the $_{F1}$ plants are backcrossed to the RP. The backcross generation (BC) so obtained and the subsequent BC progeny are backcrossed to the RP. In each BC generation, a strict selection is done for the trait/allele being introgressed from the DP because each backcrossing reduces the proportion of DP genome in the progeny to 50 % of that present in the previous generation. NIL is essentially a segment substitution version of the RP. Repeated backcrossing eliminates the DP genomic segments unlinked to the target gene and reduces the size of DP genomic region flanking the target gene due to recombination in each BC generation (Schneider 2005). NILS can be used to construct high-resolution mapping populations. Finally, they are quite useful in functional genomics; they can be used for gene expression profiling and for more direct hypothesis-driven experimentation.

Chromosomal Segment Substitution Lines: A series of homologous lines, each having a single distinct chromosome segment from a DP in the chromosome background of RP is known as . Chromosome segment substitution lines (CSSLs). The CSSLs may be produced by backcrossing the F_1 and the subsequent progeny from a cross between the DP and the RP with the RP for six generations or so, followed by self-fertilization for two or more generations to isolate lines homozygous for the introgressed segments. Selection based on markers evenly distributed over the entire genome is used to ensure that each line of the set has a distinct but slightly overlapping DP genome segment. CSSLs are a perpetual mapping resource and are suited for mapping of both oligogenes and QTLs. They can also be used for fine mapping by raising large F_2 or backcross populations following hybridization with the RP (Eshed and Zamir 1994). Evaluation of CSSLs in replicated trials over locations and years would allow the identification of such lines that have DP genomic segments with favorable effects on the traits of interest. CSSLs can be used for the detection of QTLs with small additive effects that are ordinarily masked by QTLs with larger effects in the usual mapping populations like and RILs. QTL identification using CSSLs does not require linkage map construction or statistical analysis. Further, each CSSL can be directly used for mapping and cloning of QTLs/genes and for development of elite breeding lines.

Analysis of the CSSLs PCR-based

JUST AGRICULTURE
multidisciplinary e-Newsletter

Backcross Inbred Lines: Backcross inbred lines (BILs) are developed by backcrossing the F_1 from a cross between two homozygous lines to one of the parents and continued selfing of the BC_1F_1 progeny to obtain homozygous lines. The data from BIL population were analyzed using the method for backcross F2 population and treating the heterozygotes as missing data since a method for analysis of BIL population was not available. A possible advantage of BILs may be the increased frequency of the alleles contributed by the parent used for backcrossing. Therefore, it would be desirable to use the parent with the higher value of the target trait for backcrossing with the F_1 hybrid.

Advanced Intercross Lines: An advanced intercross line (AIL) population is developed by intermating the individuals of F_2 and subsequent generations from a suitable cross. Intermating in the segregating generations maintains heterozygosity in the population and allows recombination between the QTLs and the markers linked to them in every generation leading to a more precise location of the QTLs. It was estimated that the confidence interval of QTLs would be reduced by up to five-fold in AILS as compared to that in an F_2 population. In the case of AILS, mapping resolution seems to improve for up to eight generations of intercrossing only, while it continues to improve with generation in the case of recurrent selection backcross.

Recurrent Selection Backcross Population: It refers to the population developed by backcrossing the F_1 from a cross between lines having high (DP) and low (RP) values for a quantitative trait and the subsequent generations to the RP; in each backcross generation, a predetermined number of individuals with the top phenotypic value, i.e., value close to the DP phenotype, for the trait are selected for backcrossing. RSB is proposed to be used for high-resolution QTL mapping, for which a sufficiently large number of backcrosses need to be made. Recombination between the DP QTL alleles and the linked markers will take place in each generation. Therefore, the level of heterozygosity at these marker loci will go on decreasing with the increasing number of RSB generations. In addition, in a given generation, markers located farther from the QTLs will show greater reduction in heterozygosity than those located closer to the QTLs. Thus, the frequency of heterozygosity at marker loci can be used as a criterion of localizing the QTLs.

Interconnected Mapping Populations: Interconnected mapping populations are produced by crossing a set of homozygous parental lines in such a way that two or more crosses have

JUST AGRICULTURE
multidisciplinary e-Newsletter

one parent in common. An interconnected population may consist of F_2 backcross, RIL, or DH populations generated from each of the crosses produced as per the mating design used. The usefulness of QTL findings in plant breeding depends on their general applicability and an understanding of the genetic architecture of the traits governed by the QTLs. Biparental mapping populations generate QTL information applicable to the concerned crosses, and they fail to take into account segregation of different allelic combinations of QTLs in different mapping populations and the influences of genetic background on QTL effects. Generalization of QTL findings from different biparental populations has been attempted by comparing the relative QTL positions determined from different populations by means of QTL meta-analysis, and bioinformatics tools are being developed to facilitate this analysis. In contrast, joint analysis of data from interconnected populations provides more generalized information about QTL positions and effects, increases QTL detection power, enables detection and assessment of QTL x genetic background interaction, and permits identification of markers located closer to the QTLs than do biparental populations, particularly when appropriate analysis tools are used.

Multiparent Advanced Generation Intercross Populations: The multiparent advanced generation intercross (MAGIC) populations are a collection of RILs produced from a complex cross/outbred population involving several parental lines. The parental lines may be inbred lines, clones, or individuals selected on the basis of their origin or use. MAGIC populations are perpetual, lack population structure, can be used for both linkage and association analyses, and can be developed at an appropriate stage during the intermating process to afford the desired mapping resolution. They are an ideal resource for construction of high-density maps, and they allow modeling of cytoplasmic effects. These populations can be used as training populations for genomic selection.

Nested Association Mapping Population: In order to combine the advantages of both linkage mapping and association mapping strategies, a structured population generated by crossing a set of diverse founder parents to one or two common parents has been suggested (Yu et al. 2008). Each selected founder is crossed to one or few common parents (nested parents) and a set of 250 RILs from each of these crosses is generated using the SSD method. The nested association mapping strategy enables efficient utilization of genetic and genomic resources for genetic dissection of complex traits.

JUST AGRICULTURE
multidisciplinary e-Newsletter

References

Singh BD and Singh AK (2017). Marker Assisted Plant Breeding: Principles and Practices,1st edn.Springer,New Delhi. pp 125-149.

Singh Phundan (2009). Plant Breeding : Molecular and New Approaches 2nd edn. Kalyani Publishers,New Delhi, pp 145-150.

Kover PX,Valdar W,Trakalo J (2009): A multiparent advanced generation inter cross to fine map quantitative traits in Arabidopsis thaliana. PLoS, Genet 5:e1000551