Basics of Molecular Genetic Mapping and QTL Analysis in Plants

Article · January 2009

DOI: 10.2174/978160805090110901010035

3 authors:

Jitendra Kumar
Indian Institute of Pulses Research
161 PUBLICATIONS **2,423** CITATIONS

SEE PROFILE

Aditya Pratap
Indian Institute of Pulses Research
135 PUBLICATIONS **1,856** CITATIONS

SEE PROFILE

Ramesh Solanki
Central Arid Zone Research Institute (CAZRI)
56 PUBLICATIONS **590** CITATIONS

SEE PROFILE

Basics of Molecular Genetic Mapping and QTL Analysis in Plants

Jitendra Kumar*, Aditya Pratap and RK Solanki
Indian Institute of Pulses Research, Kanpur-208024 (India)

* Corresponding author- email: jitendra73@gmail.com

Short running title: Genetic mapping and QTL analysis……..

ABSTRACT

Genetic mapping refers determination of distance and order of genes/QTL on chromosome. The conventional linkage analysis has been used to this purpose leading to publication of first genetic map of 6 sex-linked genes on a fruit fly chromosome. In plants, genetic linkge maps have been developed for major gene controlling the qualitative traits in several crops. However, possibility of locating the QTL controlling the quantitative traits has also been investigated earlier using morphological traits as genetic markers. However, the advent of molecular markers has speed up the development of dense molecular maps and QTL analysis in crop plants. This led to identification of QTL for desirable traits of agronomic importance. In this chapter, we discussed the basic principles of molecular genetic mapping and QTL analysis in plants.

Keywords: Gene, QTL, mapping population, linkage, genetic mapping, QTL analyis

INTRODUCTION

Genes or quantitative trait loci (QTL) are the genomic regions on chromosome(s), which control qualitative and quantitative variation of the trait(s). Determination of relative position between genes/QTL on a DNA molecule (chromosomes) and their distance between them is referred as genetic mapping. Classical geneticists mapped the genes controlling morphological characters and prepared the genetic maps. In 1913, Alfred Sturtevant (TH Morgan's student) [1] published first genetic map of 6 sex-linked genes on a fruit fly chromosome. Sax [2] reported possibility of locating the genomic regions controlling the quantitative traits using morphological traits as genetic markers. However, accurate and systematic genetic mapping of quantitative traits could be hampered due to the lack of a sufficient number of genetic markers covering an entire genome (dense map) in crop plants. However, advent of molecular marker technology provided various kinds of DNA-based markers, which led the development of dense molecular maps, and to date such molecular maps are available in all major food crops. Now these molecular maps are being utilized for various purposes. In this chapter, an overview of basics of molecular genetic mapping and QTL analysis was discussed.

Principle of molecular genetic mapping

Molecular markers are the DNA sequences (assay by using various molecular techniques) that have been used as genetic markers to identify the individual, species and marker-trait association. Hence, these markers similar to genes can be mapped on chromosome by using the principle of genetics involving law of independent assortment and linkage analysis. According to Mendal's second law of independent assortment, either allele of a gene/molecular marker has equal chance to go with the allele of another gene/marker located on different chromosome during the formation of gametes. Thus recombinant and parental type gametes are formed in equal frequencies i.e. a recombination frequency up to 50%. It is also hold true for genes/markers residing so apart on same chromosome (unlinked), but not for those genes that are closer together on the same chromosome. It leads to reduce frequency of recombinant type gametes less than the 50% because chance of crossing over in every cell involving in formation of gametes is reduced. It results

transmission of closely or tightly linked genes from parent to progeny in higher frequency than those genes, which are located far apart. Thus genotypic frequency of recombinant and parental genotypes are deviated from expected frequency in given mapping population (as an example of backcross mapping population is given in Figure 1). The recombination frequency between two genes is directly related to the distance between two markers/genes and determines in term of centi Morgan (1cM=1% recombinants).

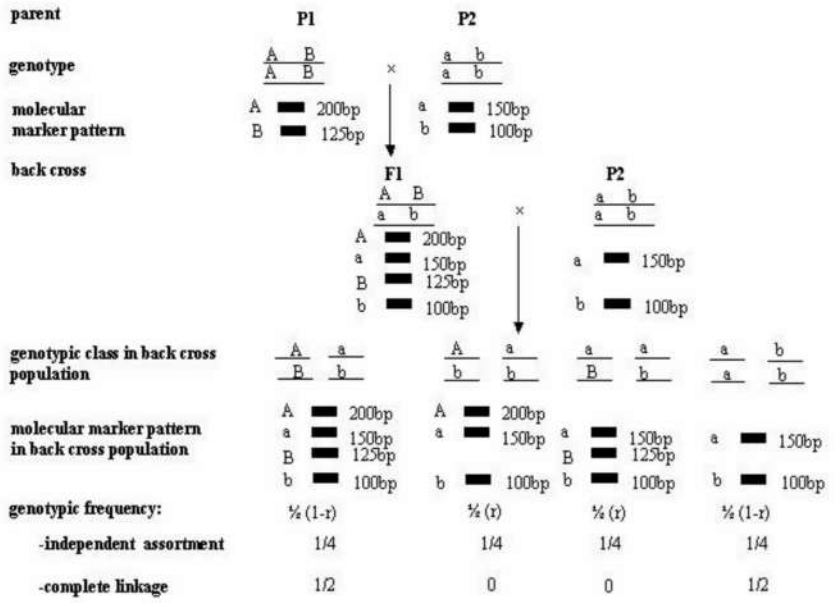

Figure 1. Expected genotypic frequency in a backcross-breeding scheme with the parents in coupling phase. The recombination value "r" can vary between 0 and 0.5, with 0 = complete linkage and 0.5 = independent assortment.

STEPS INVOLVED IN CONSTRUCTION OF MOLECULAR MAP

Above basic principles of genetics have been used for preparing the molecular genetic maps in several crops using following steps.

(i) DEVELOPMENT OF A MAPPING POPULATION

For constructing a molecular map, a mapping population is required that should be derived from the genetically diverse genotypes having contrasting genetic differences for one more traits of interest (e.g. short and tall, white and red, resistance and susceptible). These genotypes although should have sufficient polymorphism, but should not to be distantly related because it causes sterility of progenies and segregation distortion. Although it is better to use homozygous lines (inbred) for genetic mapping as used in self-pollinated species. But in case of cross-pollinated species, it is always not possible to develop such inbred lines due to inbreeding depression. Thus half or full sib progenies derived from controlled crosses can be used for genetic mapping the molecular markers. For genetic mapping in self-pollinated species, several type of mapping populations such F_2, backcross (BC), recombinant inbred lines (RILs), near-isogenic lines (NILs) and double haplod (DH) are being utilized reported in currently available literature. To decide which kind of population is used for mapping depends upon the objectives, time and nature of molecular marker used. It is critical step for successful linkage mapping. The F_2 and BC population are simplest type of mapping population that can easily be constructed in short period of time, but these populations provide spurious linkage due to their derivation after only one or two round of recombination events and are also to be mortal. But on other hand use of RILs and NILs populations derived from several round of recombination events and hence reduces chance to get the spurious linkage, are best for genetic mapping, although it required several years to develop. In addition to this RIL, NILs and DH are immortal population can be used one laboratory to another laboratories. Staub and Serquen [3] reviewed the genetic information for different types of mapping populations in relation to dominant vs co-dominant markers. Map constriction using co-dominant markers is more accurate than the dominant markers. Less information is obtained in BC population because only one recombinant gamete is sampled per plant.

(ii) SIZE OF MAPPING POPULATION

For developing an accurate genetic map, it is essential to use the sufficient number of individuals for genetic mapping. The studies suggest the use of 50 to 1000 individuals of either population (F_2, BC or DH), but use of lowest number of individuals provided

several fragmented linkage groups and inaccurate locus order. Use of the larger the population size provides more accurate of genetic map because it increases possibility of getting recombinants in a population. Most of genetic maps were constructed in crop plant used a population size of 50 to 100 individuals, but for an accurate genetic map, a population of 200 individual is shown reasonable using either type of populations. However for identification of tightly linked molecular markers, a population of more than 1000 individuals has to be screened for genetic mapping in several crop species.

(iii) SELECTION OF MOLECULAR MARKERS SYSTEMS

During past two decades, a number of molecular markers have been developed including restriction fragment length polymorphisms (RFLP), random amplified polymorphic DNAs (RAPDs), amplified fragment length polymorphisms (AFLPs), simple sequence repeat (SSR), single nicleotide polymorphisms (SNPs) and diversity arrays technology (DArT). The classification and general features of these molecular markers have been reviewed extensively [4, 5].

RFLP is first molecular marker system that have been used extensively in preparation of molecular maps in crop plants. This is due to the co-dominant nature, reproducibility one lab to another and locus specificity that allow synteny studies. However, further application of this marker system in crop improvement have several problems including hazardous to health, required high quality of DNA, requires development of specific probes, low polymorphism, time consuming and laborious.

As the advent of polymerase chain reaction, new marker system have been developed which have become popular for past years. Among the PCR-based molecular markers, microsatellite or simple sequence repeat (SSR) markers have been shown marker of choice for construction of molecular maps. The main advantages of SSR markers are high level of polymorphism, high reproducibility, co-dominant nature, locus specificity and easy in use. However, development of these molecular requires the genomic sequence that makes them costly and has become limited to few major crop species. However, the transferable SSR markers increase the possibility of using these markers in closely related species. It is particularly shown for EST derived SSR markers because these markers are developed from transcribed regions of the genome, which are

more conserved over the genera. Furthermore, these markers have been shown more useful to identify the candidate genes. Although PCR-based RAPD markers have been used to construct linkage maps in several species without the prior knowledge of genomic sequence information, poor reproducibility of these markers have restricted wide acceptance in genetic mapping.

Amplified fragment length polymorphism (AFLP) is another widely used DNA markers for genetic mapping especially to increase the density of molecular maps. These markers are highly polymorphic, reproducible, and can be used for any organism without initial investment in primer/probe development and sequence information. They generally provide good coverage, but clustering of AFPL markers especially from the methylation insensitive EcoRI-enzyme may aggravate clustering in the highly methylated centromere and surrounding heterochromatin regions of the genome.

Diversity arrays technology (DArT) is a microarrays-based molecular marker system used in genetic mapping and fingerprinting in several crops including wheat [6, 7] , *Arabidopsis* [8] and barley [9]. The beauty of this marker system is that it can investigate several marker loci simultaneously and hence it is one of high throughput marker system. In addition to this, high polymorphic nature and high reproducibility make them further marker of choice, but dominant inheritance is still a limitation for mapping.

The variation of a single base in the DNA sequences of two individuals has been used as genetic markers. These differences among the individuals are known as single nucleotide polymorphism (SNP). Generally, SNPs are to be bi-allelic. It means that two individuals differing at a single base will have either of bases (A/T or C/G). But they become multiallelic when assess at the level of haplotype (a group of linked SNPs in a 500-1000bp DNA) and hence it makes them highly polymorphic. SNPs may be either dominant or co-dominate depending upon techniques used to assess them. For last few years, a rapid progress has been shown in mapping of SNPs and associated with a number of agronomically important traits. However detection of SNPs demands extensive investment for highly specialized equipment and skilled manpower.

(iv) GENOTYPING OF MAPPING POPULATION

A selected molecular marker system is first used to screen the parental genotypes involved in mapping population in order to identify the sufficient polymorphic markers for molecular map construction. In general, cross-pollinated species have higher level of polymorphism compared to self-pollinated species. A number of polymorphic markers identified are then used to genotype the all individuals of a mapping population. Genotyping data of each individual is scored as 1 and 0. However, it depends upon the molecular markers and type of mapping populations used for constructing the molecular maps. For example, if a RIL population is genotyped with SSR markers, genotyping of individuals are done as 1 (Parent 1 allele) and 0 (Parent 2 allele). But in case of F_2 population, heterozygotes also present in the population and hence genotyping data is scored as 1 (Parent 1 allele), 2 (F_1; both allele) and 0 (Parent 2 allele). The dominant markers cannot differentiate the dominant homozygotes from heterozygotes individuals and then F_3 population can be used to differentiate them. A good genetic map not only depends upon a marker system revealing high level of polymorphism, but also depends upon genotyping data used to construct it. Therefore the data must be critically checked for all possible errors, such as typological error, missing data, genotyping coding error, order of genotypes along all loci, etc. It is absolutely essential that the order of the individuals is identical over all loci in the data file.

(v) PERFORM THE LINKAGE ANALYSIS AND MAP CONSTRUCTION

Genotyping data recoded on a mapping population is used to prepare a molecular map on the basis of linkage between the markers using one of the following computer software packages.

MapMaker/Exp (ftp://genome.wi.mit.edu/pub/mapmaker3/): freely distributed, analyzes F2 and BC mapping populations [10].

Mendel: http://gnome.agrenv.mcgill.ca/info/gmendel.htm): freely distributed, analyzes all types of mapping populations; can combine maps of different mapping populations provided there are common markers [11].

JoinMap (http://www.cpro.dlo.nl.cbw/): analyzes all types of mapping populations; can combine maps of different mapping populations provided there are common markers [12].

LINKAGE-1: It is capable of analyzing in a single run an unlimited number of progenies generated from a variety of genetic situations. These families can represent F2 and backcross types as well as all other combinations of the allowable single- factor segregation ratios (e.g., where one locus is in testcross mode and the other has a 1:2:1 expected ratio). The program accepts both dominant (3:1 and 1:1 expected ratios) and co-dominant (1:1, 1:2:1, and 1:1:1:1 expected ratios) genes as segregating loci. Input for each family consists of single-individual genotype data for each segregating locus. Program is available with author and can be used on request [13].

MAP Manager QTX: It includes functions for mapping both Mendelian and quantitative trait loci and it is an enhanced version of Map Manager QT. It currently is available at http://mapmgr.roswellpark.org/mmQTX.html. [14].

The basic principles used to construct the molecular map in these programs are similar as described below:

(a) TEST FOR SEGREGATION DISTORTION

Molecular markers used to genotype a population must be follow the law of segregation for constructing a precise molecular maps and each marker should segregate either into 1:1 ratio in DH, RILs and BC population or 1:2:1 segregation in F_2 population. But sometimes a deviation between observed and expected genotypic frequency of molecular markers called segregation distortion are observed in a given population, which violates the law of segregation and renders conventional genetic theory and analysis of linkage invalid [15]. Segregation distortion was first reported in maize by Mangelsdorf and Jones [16] on the basis of linkage between the gametophyte factor *Ga1* and the *Su* allele for starchy endosperm. Thereafter it has been reported in several other plant species [17]. However extend of segregation distortion found variable among different mapping populations. For example, double haploid (DH) and recombinant inbred lines (RIL) populations usually have extreme segregation distortion. In DH populations, many recessive lethal genes become homozygous and express, leading to a high percentage of markers showing segregation distortion as reported in coffee [18]. Similar results also obtained in DH and BC1 population of cotton [19], although backcross populations usually have relatively fewer segregation distortions [20]. The high segregation distortion of the RIL population was inferred to be related to environmental and artificial selection

over several generations during its development (Wang et al. 2003). Significant segregation distortion also reported in F_2 population [21]. Many factors can cause the segregation distortion including errors in marker genotyping and statistical analysis, residual heterozygosity in parental lines, a mutation within the binding site for primers, pollen abortion, pollen tube competition, competitive fertilization, zygotic selection, environmental factors, heterogeneous recombination, gene diversion, transgenic silencing, selection pressure during *in vitro* androgenesis and chromosome instability resulting from aneuploid and unstable translocation [17].

Segregation distortion influences genetic distance between markers and the order of markers on linkage groups [22, 23]. Because it estimates recombination fraction significantly either higher or lower in a mapping population, as reported in coffee and brassica [17], which consequently affects the accuracy of a genetic map. Therefore, molecular markers showing segregation distortion should be excluded during preparation of map, although this treatment usually reduces coverage of the genome and qualitative or quantitative trait loci might be missed. However, such makers can be used for mapping with normally segregated markers, the fate of the distorted markers should be determined on basis of how much the map was affected. But there is not a proper standard for deleting the distorted markers from a genetic map. Effects of segregation distortion on linkage were estimated with independent χ^2 test and LOD values, and the use of a more stringent linkage test was recommended for markers showing extreme segregation distortion, such as higher LOD values or lower maximum linkage distance [24, 25]. Comparative mapping among different types of populations from the same cross or among different populations of the same type can be one of the ways to identify spurious linkage, as done in maize to identify loose and spurious linkages [24, 26, 27]. A methods simulated population sampling criteria' (SPSC) also developed to detect the extent of segregation distortion of family lines and markers and to adjust the RIL population [28]. Besides Lorieux and co-workers developed computer software, Mapdisto, for genetic map construction with segregation data including distorted markers. After several improvements, this software can now solve map construction and map drawing and deal with F_2, BC_1, DH, and single seed descent (SSD) populations. Using this software,

Lashermes *et al.* [18] corrected results of two linkage groups containing distorted markers in coffee [17].

(b) ESTABLISHING LINKAGE GROUPS, MAP DISTANCE AND LOCUS ORDER

Molecular markers then are assigned to linkage groups on the basis of a logarithm of odds (LOD) value. LOD refers to the ratio of the probability that two loci are linked with a given recombination value over a probability that the two are not linked [12, 29]. The critical LOD scores used to establish linkage groups and calculate map distances are called 'linklod' and 'maplod', respectively [30, 31]. Marker pairs with a recombination LOD score above a critical 'linklod' are considered to be linked whereas those with a LOD score less than 'linklod' are considered unlinked. Several researchers used a 'linklod' value of 3 as the minimum threshold value in order to decide whether or not loci were linked. A LOD value of 3 between two markers indicates that linkage is 1000 times more likely than no linkage [12]. Although it has also shown with modern data sets with many markers, especially those species having large numbers of chromosomes that even using a 'linklod' of > 6 may lead to false positive linkage [30]. Thus various options (e.g., changing the parameters of analyses, excluding loci or individuals, generation of additional marker data for linkage groups with few number of markers, etc.) can be tested until in order to establish satisfactory linkage groups equal to haploid chromosome number. It is necessary because higher critical LOD values will result in more number of fragmented linkage groups while small LOD values will tend to create few linkage groups with large number of markers per group. In practice, determining number of linkage groups is usually not a straightforward task because; a) loci on different chromosomes may appear to be linked by chance (spurious linkage). b) two or more linkage groups can be obtained for each chromosome, which results to the total number of linkage groups much higher than the haploid chromosome numbers.

For determining the map distance and locus order, several parameters including a recombination threshold value, minimum 'maplod', jump threshold value have to specify a LOD score above 'maplod' is only used in the calculation of map distances, of which choice is arbitrary and can be varied from 0.01 to 3.0. A 'maplod' equal to 0.01 indicates very weak linkage. Computer programs used to prepare molecular maps use either

Haldane or Kosambi mapping functions, which translates recombination frequencies into map distances and vice versa. The map distance calculated by Haldane's mapping function [32] is grater than Kosambi's mapping function [33] because it assumes no interference between crossovers in meiosis, while later one assumes a certain degree of interference. The statistical programs perform locus ordering using one of the locus-ordering criteria: weighted least squares, maximum likelihood and minimum sum of adjacent recombination fractions. Hackett and Broadfoot [34] performed a simulation study to compare the performance of these three locus-ordering criteria in the presence of missing values, typing errors and distorted segregation ratios. The authors concluded that map inflation was more extreme using the maximum likelihood criterion than using weighted least squares. The advantage of the weighted least-squares approach is that the distances between markers are calculated from the map distances

QTL ANALYSIS

The most important application of molecular markers is locate the genomic regions on chromosome that affect phenotypic expression of a desirable trait. Analysis of such genomic regions can be done with or without molecular maps, which discussed below. An outline of genetic mapping of genes/QTL is given in Figure 2.

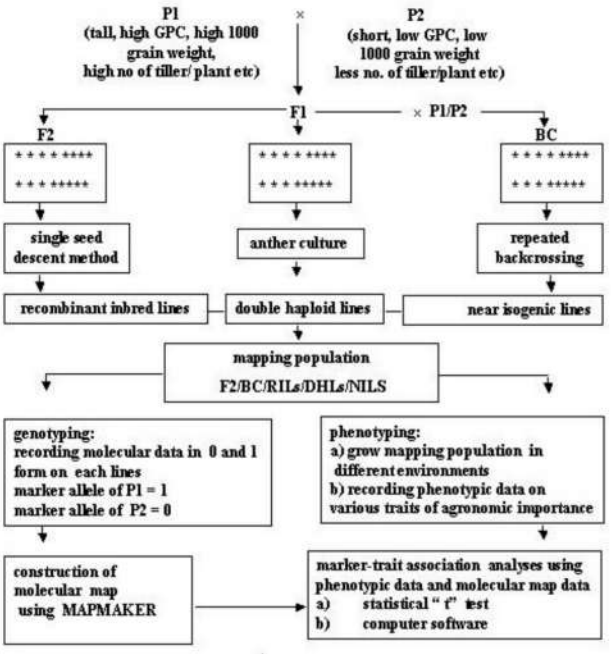

Figure 2. An outline of genetic mapping of genes/QTL in plants

(i) QTL ANALYSIS WITHOUT USING MOLECULAR MAP

Thoday [35] gave the basic idea of QTL analysis. According to him, if genetic markers are scattered throughout the genome of an organism of interest, the segregation of these markers can be used to direct and estimate the effects of linked QTL, making possible the mapping and characterization of underlying QTL. The QTL method involves searching for associations between the molecular marker and the trait of interest in a segregating population such as F_2, backcross (BC), recombinant inbred lines (RILs), and double haploid lines (DHLs). These association although can be identified through by several traditional methods like single locus analysis, AMOVA, t-test and regression analysis. The traditional method to detect a QTL in the vicinity of a marker is studying single genetic markers one at a time. In these methods, the phenotypic means for progeny of each marker class are compared (e.g., means of the marker classes AA, Aa, aa). The difference between two means provides an estimate of the phenotypic effect of substituting an "A" allele by an "a" allele at the QTL. To test whether the inferred phenotypic effect is significantly different from zero, a simple statistical test, such as t-test or F-test, is used. A significant value indicates that a QTL is located in the vicinity of the marker. Single point analysis does not require a complete molecular linkage map. The further a QTL is from the marker, the less likely it is to be detected statistically due to crossover events between the marker and the gene.

(ii) QTL INTERVAL MAPPING BY USING MOLECULAR MAPS OR BY PREPARING FRESH MOLECULAR MAPS

Because basic QTL analysis approach has several limitations and hence scientists have developed new strategies called interval mapping for QTL analysis that involves use of molecular maps. QTL interval mapping is probably the most common method of QTL analysis. A well-known example is the Mapmaker/QTL developed by Lincoln *et al.* [36]. The principal behind interval mapping is to test a model for the presence of a QTL at many positions between two mapped marker loci. Interval mapping includes simple

interval mapping (SIM) and composite interval mapping. In simple interval mapping, an individual interval is scanned at a time for the presence of QTL. However, in this case the effects of additional QTL can contribute to sampling variance and when two-linked QTL can cause biased estimated due to their combined effects. In order to overcome these limitations of SIM, composite interval mapping (CIM) was proposed as a method of solution [37, 38, 39]. CIM will perform the analysis in the usual way, except that the variance from other QTL is accounted for by including partial regression coefficients from markers ("cofactors") in other regions of the genome. CIM gives more power and precision than simple interval mapping (SIM) because the effects of other QTL are not present as residual variance. CIM can remove the bias that can be caused by QTL that are linked to the position being tested. Following two approaches are used in interval mapping

a) MAXIMUM LIKELIHOOD APPROACH: If it is assumed that a QTL is located between two markers, the 2-loci marker genotypes (i.e., AABB, AAbb, aaBB, aabb for DH progeny) each contain mixtures of QTL genotypes. Maximum likelihood involves searching for QTL parameters that give the best approximation for quantitative trait distributions that are observed for each marker class. Models are evaluated by computing the likelihood of the observed distributions with and without fitting a QTL effect. The map position of a QTL is determined as the maximum likelihood from the distribution of likelihood values (LOD scores: ratio of likelihood that the effect occurs by linkage: likelihood that the effect occurs by chance), calculated for each locus

b) INTERVAL MAPPING BY REGRESSION: Interval mapping by regression [40] was developed primarily as a simplification of the maximum likelihood method. It is essentially the same as the method of basic QTL analysis (regression on coded marker genotypes) except that phenotypes are regressed on QTL genotypes. Since the QTL genotypes are unknown, they are replaced by probabilities estimated from the nearest flanking markers. In most cases, regression mapping gives estimates of QTL position and effect that are almost identical to those given by the maximum likelihood method. The approximation deviates only at places where there are large gaps, or many missing genotypes.

To date there are over 100 genetic analysis software packages (linkage analysis and QTL mapping), but here we listed some features of the most commonly used software packages.

MapMaker/QTL (ftp://genome.wi.mit.edu/pub/mapmaker3/) is the original QTL mapping software for IBM computer. It is user-friendly, freely distributed, and runs on almost all platforms. It will analyze F_2 or backcross data using standard interval mapping.

MQTL: is an IBM computer program for composite interval mapping in multiple environments. It can also perform simple interval mapping. Currently, MQTL is restricted to the analysis of data from homozygous progeny (double haploids, or recombinant inbred lines). Progeny types with more than two marker classes (e.g., F_2) are not handled.

PLABQTL (http://www.uni-hohenheim.de/~ipspwww/soft.html) is a freely distributed IBM computer program for composite interval mapping and simple interval mapping of **QTL** : Its main purpose is to localize and characterize QTL in mapping populations derived from a biparental cross by selfing or production of double haploids. Currently, this program is the easiest software for composite interval mapping.

QTL Cartographer (http://statgen.mcsu.edu/qtlcart/cartographer.html) is QTL software written for either UNIX, Macintosh, or Windows. It performs single-marker regression, interval mapping, and composite interval mapping. It permits analysis from F_2 or backcross populations. It displays map positions of QTL using the GNUPLOT software.

MapQTL (http://www.cpro.dlo.nl/cbw/)

Qgene is a QTL mapping and marker-aided breeding package written for Macintosh. It has a user-friendly graphical interface and produces graphical outputs. QTL mapping is conducted by either single-marker regression or interval regression.

SAS is general statistical analysis software. It can detect QTL by identifying associations between marker genotype and quantitative trait phenotype by single marker analysis approach such as ANOVA, t-test, GLM or REG.

OTHER APPLICATIONS OF MOLECULAR GENETIC MAPPING

i) Introgression of desirable genes or QTLs through marker-assisted selection.

ii) Allow comparative mapping between different species in order to evaluate similarity between genes orders and function in the expression of a phenotype.

iii) Provide a framework for anchoring with physical maps based on chromosome translocations, DNA sequence or other direct measures.

iv) Constitute the first step towards positional or map-based cloning of genes responsible for economically important traits.

References

[1] Sturtevant AH. The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. Journal of Experimental Zoology 1913; 14: 43-59

[2] Sax K. The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. Genetics 1923; 8: 552–560

[3] Staub JE, Serquen FC. Genetic markers, map construction, and their application in plant breeding. HortScience 1996; 31: 729-740

[4] Gupta PK, Varshney RK, Sharma PC, Ramesh B. Molecular markers and their application in wheat breeding: a review. Plant Breed 1999; 118: 369-390

[5] Khlestkina EK, Salina EA. SNP Markers: Methods of analysis, ways of development, and comparison on an example of common wheat. Russian Journal of Genetics 2006; 42: 585–594

[6] Akbari M, Wenzl P, Caig V, Carlig J, Xia L, Yang S, Uszynski G, Mohler V, Ehmensiek A, Howes N, Sharp P, Huttner E, Kilian A. Diversity arrays technology (DArT) for high throughput profiling of the hexaploid wheat genome. Theor Appl Genet 2006; 113:1409-1420

[7] Semagn K, Bjørnstad Å, Ndjiondjop MN. Principles, requirements and prospects of genetic mapping in plants. African Journal of Biotechnology 2006; 5: 2569-2587

[8] Wittenberg AHJ, van der Lee T, Cayla C, Kilian A, Visser RGF, Schouten HJ. Validation of the high-throughput marker technology DArT using the model plant *Arabidopsis thaliana*. Mol Genet Genomics 2005; 274: 30-39

[9] Wenzl P, Carling J, Kudrna D, Jaccoud D, Huttner E, Kleinhofs A, Kilian A. Diversity Arrays Technology (DArT) for whole-genome profiling of barley. Proc Natl Acad Sci USA 2004; 101: 9915-9920

[10] Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1987; 1: 174–181

[11] Echt C, Knapp S, Liu BH. Genome mapping with non-inbred crosses using Gmendel 2.0. Maize Genet. Coop. Newsletter 1992; 66: 27-29

[12] Stam P Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J 1993a; 3: 739-744

[13] Suiter KA, Wendel JF, Case JS. Linkage-I: A pascal computer program for the detection and analysis of genetic linkage. J Hered 1983; 74: 203-204

[14] Manly KF, Cudmore Jr RH, Meer JM Map Manager QTX, cross platform software for genetic mapping. Mamm. Genome 2001; 12: 930–932

[15] Lu H, Romero-Severson J, Bernardo R. Chromosomal regions associated with segregation distortion in maize. Theor Appl Genet 2002; 105: 622-628

[16] Mangelsdorf PC, Jones DF The expression of Mendelian factors in the gametophyte of maize. Genetics 1926; 11: 423–455

[17] Song XL, Sun XZ, Zhang TZ. Segregation distortion and its effect on genetic mapping in plants. Chinese Journal of Agricultural Biotechnology 2006; 3: 163–169

[18] Lashermes P, Combes MC, Prakash NS, Trouslot P, Lorieux M and Charrier A. Genetic linkage map of *Coffea canephora*: effect of segregation distortion and analysis of recombination rate in male and female meiosis. Genome 2001; 44: 589–596

[19] Song XL, Wang K, Guo WZ, Zhang J, Zhang TZ. A comparison of genetic maps constructed from haploid and BC_1 mapping populations from the same crossing between *Gossypium hirsutum* L.rG. *barbadense* L. Genome 2005; 48: 378–39

[20] Kianian SF, Quiros CF Generation of a *Brassica oleracea* composite RFLP map: linkage arrangements among various populations and evolutionary implications. Theor Appl Genet 1992; 84: 544–554

[21] Wang YJ, Wu XL, He CY, Zhang JS, Chen S and Gai JY. A soybean genetic map constructed after the population being tested and adjusted. Scientia Agricultica Sinica 2003; 11: 1254–1260

[22] Lorieux M, Goffinet B, Perrier X, Gonzalez D de L, Lanaud C. Maximum-likelihood models for mapping genetic markers showing segregation distortion. 1. Backcross populations. Theor Appl Genet 1995a; 90: 73–80

[23] Lorieux M, Perrier X, Goffinet B, Lanaud C, Gonzalez D de L. Maximum-likelihood models for mapping genetic markers showing segregation distortion. 2. F2 populations. Theor Appl Genet 1995b; 90: 81–89

[24] Murigneux A, Baud S, Beckert M. Molecular and morphological evaluation of doubled-haploid lines in maize. 2. Comparison with single-seed descent lines. Theor Appl Genet 1993; 87: 278–287

[25] Manly KF Establishing genetic linkage using recombinant inbred lines with an abnormal segregation ratio. Genetics 1994; 136: 1433–1434

[26] Bentolila S, Hardy T, Guitton C, Freyssinet G. Comparative genetic analyses of F2 plants and anther culture-derived plants of maize. Genome 1992; 35: 575–582

[27] Cloutier S, Cappadocia M, Landry BS. Analysis of RFLP mapping inaccuracy in *Brassica napus* L. Genetics 1997; 95: 83–91.

[28] Wang YJ, Wu XL, Yu DY, Zhang YM, Chen SY, Gai JY. Method of evaluation and adjustment of recombinant inbred line population and its application to the soybean RIL population NJRIKY. Acta Agronomica Sinica 2004; 30: 413–418

[29] Risch N. Genetic linkage: Interpreting LOD scores. Science 1992; 255: 803–804

[30] Stam P. JoinMap Version 1.4: A computer program to generate genetic linkage maps. CPRO-DLO, Wageningen. Centre for Plant Breeding and Reproduction Research CPRO-DLO, Wageningen, The Netherlands; 1993b.

[31] Ortiz JPA, Pessino SC, Bhat V, Hayward MD, Quarin CL. A genetic linkage map of diploid *Paspalum notatum*. Crop Sci 2001; 41: 823-830

[32] Haldane JBS. The cytological basis of genetical interference. Cytologia 1931; 3:54–65

[33] Kosambi DD. The estimation of map distances from recombination values. Ann Eugen 1944; 12: 172–175

[34] Hackett CA, Broadfoot LB. Effects of genotyping errors, missing values and segregation distortion in molecular marker data on the construction of linkage maps. Heredity 2003; 90: 33-38

[35] Thoday JM. Location of polygenes. Nature 1961; 191: 368–370

[36] Lincoln SE, Daly MJ, Lander ES. Constructing genetic linkage maps with MAPMAKER/EXP version 3.0: A tutorial and reference manual. Whitehead Institute for Biomedical Research, Cambridge, MA; 1993.

[37] Jansen RC, Stam P. High resolution of quantitative traits into multiple loci via interval mapping. Genetics 1994; 136: 1447–1455

[38] Utz HF, Melchinger AE. Comparison of different approaches to interval mapping of quantitative trait loci. In: van Ooijen JW, Jansen J (Eds.) Biometrics in Plant Breeding: Applications of Molecular Markers. Proceedings of the Ninth Meeting of the EUCARPIA Section Biometrics in Plant Breeding, 6 – 8 July 1994, Wageningen. CPRO-DLO, Wageningen, The Netherlands 1994; pp. 195–204

[39] Zeng ZB. Precision mapping of quantitative trait loci. Genetics 1994; 136: 1457-1468

[40] Haley CS, Knott SA. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity 1992; 69: 315-324.