Research Review

© 2023, Centre for Advanced Research in Agricultural Sciences
Research Journal of Agricultural Sciences **14**(6): 1888–1891 (Nov-Dec 2023)
DOI: http://rjas.org/Article/Article/5704

ISSN: 0976-1675 (P)
ISSN: 2249-4538 (E)

The Principles of Quantitative Trait Loci (QTL) Mapping

M. SRIDHAR REDDY*[1] and P. OSMAN BASHA[2]

[1] Department of Environmental Sciences, Yogi Vemana University, Kadapa - 516 005, Andhra Pradesh, India
[2] Department of Genetics and Genomics, Yogi Vemana University, Kadapa - 516 005, Andhra Pradesh, India

Received: 30 Jun 2023; Revised accepted: 18 Nov 2023; Published online: 11 Dec 2023

Abstract

The objective of this review paper is to emphasize the significance of Quantitative Trait Loci (QTL) mapping and the driving concepts behind these analyses. Molecular marker and trait data analysis together has emerged as a crucial tool for biologists to analyze the genetics of complex traits. Theoretically, a QTL can refer to a single gene or a set of linked genes or an area of the genome that influences a trait. QTL mapping is a marker and trait association-based technique that can be used successfully for gene pyramiding, screening of germplasm for abiotic and biotic stresses. Often users have been deterred from understanding what the procedures adapted and understanding different approaches merits and limitations available in mapping due to the mathematical and statistical methodologies used. Different mapping populations, such as F_2, back crosses, recombinant inbred lines, and double haploid lines, are crucial for trait data processing. A crucial characteristic of this kind of population is strong linkage disequilibrium at marker loci and alleles of linked loci which influencing the trait. The co-segregation of marker loci and QTL is the fundamental principle. The single marker approach (SMA), simple interval mapping (SIM), composite interval mapping (CIM), and multiple interval mapping (MIM) are methods developed for QTL mapping.

***Key words*:** Quantitative trait loci, DNA markers, Mapping population, QTL mapping

Quantitative Trait Loci (QTL) mapping is a genome-wide inference of the association between different chromosome locations and characteristics for a set of quantitative traits in terms of the number, chromosomal locations, and their effects. Traits are classified into two types; they are quantitative traits and qualitative traits. In this particular instance, quantitative variation is continuous, whereas qualitative variation is discontinuous. Qualitative variation is typically governed by number of genes and placed into a few distinct phenotypic classes known as discrete classes. Individual genotypes can be predicted using these classes. Molecular markers are perfect for investigating and mapping QTLs, which can then be used effectively in Molecular Assisted Selection. Angaji [1] defined QTL mapping is defined as the marker-facilitated genetic dissection of complex variations via suitable experimental design and statistical segregation studies. The analysis is based on calculations of mean difference among lines with distinctive marker alleles. This method is the first step in identifying the desired target genes for marker-assisted selection. Only genes associated with disease resistance and stress tolerance have shown to be effective thus far. QTL mapping is a fundamental research activity that requires meticulous cross-planning and very accurate phenotyping. The invention of genetic (or molecular) markers in the 1980s marked a significant advance in the identification of quantitative traits that provided chances to select for QTLs [2].

Why QTL mapping

The application of QTL mapping provides a straightforward method for determining the number of genes influencing a characteristic, the location of the genes, and the impact of the dosage of these genes on variation of trait. The initial stage in map-based cloning is genetic mapping. It is employed in DNA-based marker assisted selection (MAS) and the investigation of the linkages between particular genes. Finding QTLs and determining their locations and effective sizes for genes with minor effects are challenging tasks. Generally, for locating single genes with significant effects, QTL mapping works best. The reliability of the QTL map will be compromised by low quality QTL phenotyping experiment data collection. It requires a highly heritable phenotypic screening system to get reliable QTL's. Finding the area of the genome or gene that controls the quantitative trait of interest and analyzing the QTL's influence on the trait are the goals of QTL mapping [2]. The questions that are typically raised during QTL analysis are as follows: i. To what extent does a particular location contribute to variance in the target trait of interest? ii. Which type of gene action additional or dominant is related to QTL. iii. What kind of alleles are associated with a favorable effect?

Requirements of QTL mapping

The chosen parents should typically differ from one other in detectable ways, and the mapping populations resulting from these crossings may contain lines that are having gradient of resistant to susceptible to disease. It is preferable to use highly trustworthy screening methods that can tell resistant from susceptible lines apart. Instead of using data from a single trial, analysis should be based on the means of multiple trials. Make sure the collected data repeatability is as high as feasible

***Correspondence to:** P. Osman Basha, E-mail: osmanbasha@yahoo.co.in; Tel: +91 9701725190

Citation: Reddy MS, Basha PO. 2023. The principles of quantitative trait loci (QTL) mapping. *Res. Jr. Agril. Sci.* 14(6): 1888-1891.

(0.7 or higher). In another words, A mapping population developed from morphologically different parents, a saturated genetic linkage maps developed based on molecular markers, consistent phenotypic evaluation of mapping population, and selection of suitable statistical tool to analyze the genotypic data along with phenotypic data are the fundamental prerequisites for QTL mapping.

QTL mapping strategies

The same basic methodology is used in all marker-based mapping studies: (i). Selection of parent's distinct phenotype from one another. (ii). Selection of polymorphic markers among the parents. (iii). Development of F_2, BC (back crosses), RIL (recombinant inbred lines), NIL (near isogenic lines) and DH (double haploid) mapping populations. (iv). Phenotype evaluation. (v). Compare the means of the "AA" (parent 1 phenotype) and "aa" (parent 2 phenotype) lines at each marker locus. (vi). There is a greater likelihood of detecting QTLs if there is a larger difference between the means of the MM and mm lines.

Principle of QTL mapping

Marker loci and quantitative traits co-segregate with each other from one generation to next generation, according to the core idea. Co segregation occurs as a result of linkage between markers and quantitative traits. To determine the linkage, progeny testing is used to divide the mapping population into different genotypic classes. There should be a strong association between the molecular marker and the target gene locus on chromosome so that they will co-segregate in mapping population. "Linkage disequilibrium" describes situations in which genes fail to segregate independently. Thereby, QTL analysis is reliant on linkage disequilibrium to assess whether the QTL is linked to a marker or not [2].

Factors effecting QTL mapping

QTLs are statistically analyzed and predicted information determined based on the data generated experimentally and the efficiency of QTL mapping is influenced by the number of genes controlling the assessed traits, heritability of the genes in segregating mapping population, type and size of mapping population, type and number of markers used in linkage map construction and phenotyping of mapping population variables. The success of QTL mapping depends on the gene's location on the chromosome in relation to polymorphic markers. QTLs will have a better chance of being discovered if they remain near to the relevant genetic marker. There will be a greater likelihood of crossing over if genes are present away from the concerned genetic marker.

The detection and analysis of crossing overs between QTL and molecular markers were made based on genetic molecular markers bands pattern. It will be tricky to pinpoint the location of the target loci if the distance is great. In general, characters controlled by single genes or few genes typically have higher heritability than those controlled by polygene. For QTL mapping, F_2, BC, and RIL mapping population is necessary and development of these population requires through understanding of segregation analysis.

The discovery of QTLs was substantially influenced by the size of the mapping population used in QTL mapping. Development and selection of large mapping population make it possible to identified QTL having significant impact on traits. When small mapping samples are chosen, QTL with small effects cannot be seen, but there is a chance to find QTL with large impacts on variables. The type of molecular markers used to build genetic linkage maps had a big impact on QTLs detection. Additionally, the accuracy of estimating both the QTL and its effect will increase with the usage of more markers. Co-dominant molecular markers typically exhibit three types of genetic variation by detecting the heterozygous genetic condition in mapping population. In contrast, dominant markers typically show two genetic variations and cannot be detected individuals with heterozygous condition in the mapping population. Hence, co-dominant markers reveal more information about chromosomal recombination events than dominant markers [2].

Furthermore, the mapping population's phenotyping is one of the important factors in identifying the QTLs. A limited number of incomplete or absent data points can be accepted as long as the target quantitative characteristics are determined as accurately as feasible. The mapping population sample size and then the genome-wide coverage of molecular markers serve as elements on the ability to resolve the QTL site. Even though mapping populations often have greater sample sizes, sometimes missing data or skewed allele frequencies cause the sample's real size to be less, which further hinders statistical analysis. Population size must occasionally be compromised for data quality, and as a result, only major QTL can be identified. Typically, QTL data is combined across replications collected from different locations to provide a single quantitative trait for each individual of population to have a improved understanding of QTL and environment interaction. Therefore, whenever an allelic substitution's orientation can be shown, QTL data may demonstrate that directional selection is present. This method was employed to quantify the dominant selective parameters accountable for evolution of such dissimilar organisms as sunflowers and Lake Malawi cichlids [3-4].

Quantitative trait loci detection

Most of the important traits including yield, quality, days taken to flower, biotic and abiotic stress tolerance and disease resistance are controlled by many genes with small effects called QTLs. The availability of different DNA molecular markers in the recent years has led to considerable progress in QTLs and gene mapping in plants [5-6]. The molecular markers have been extensively used for construction of genetic maps utilizing mapping populations, which could consist of an F_2, a backcross, doubled haploids (DHs) and recombinant inbred lines (RILs) [7]. The (Fig 1-2) illustrates the design of recombinant inbred lines (RIL) and near-isogenic line (NIL) population for the detection of QTLs with a high likelihood. When a donor and recurrent parent are crossed initially in NIL population development, the donor genome contribution is reduced as a result of successive back crossings to the recurrent parent. The construction of an array of NILs that cover the genome is possible with marker assisted selection (MAS). For the purpose of identifying and locating QTLs, the generated population can be examined for a variety of phenotypic features. Several researchers have development of recombinant inbreed lines to map QTLs. The RIL population was developed through multiple rounds of self-crossing. In this population, the selection of specific, homozygous, single, overlapping chromosome segments simplifies QTL localization and identifies associated genetic markers for crop improvement [8].

Molecular markers

A molecular marker is defined as a DNA fragment found at a specific location in the genome. The characteristics of molecular markers, like alleles, are different to each individual. As a result, the distances between linked molecular markers can be calculated using recombination in the segregating population. The molecular markers are developed based on polymorphism in DNA sequences. Diversity analysis studies

have been carried out using different marker series. Different types of markers developed as based on Southern hybridization, include restriction fragment length polymorphisms (RFLPs), which were the first DNA based markers, that were used in various crops [9-12]. The PCR based markers include simple sequence repeats (SSRs), random amplified polymorphic DNA (RAPDs), sequence tagged sites (STSs), sequence characterized amplified regions (SCARs), inter-simple sequence repeat amplification (ISSR), cleaved amplified length polymorphic sequences (CAPs), amplified fragment length polymorphisms (AFLPs), DNA amplification fingerprinting (DAF) and single nucleotide polymorphism (SNP) [13-19]. Out of these, simple sequence repeats (SSRs) or microsatellites and restriction fragment length polymorphisms (RFLPs) molecular markers were extensively used for QTL mapping due to robust and reliability.

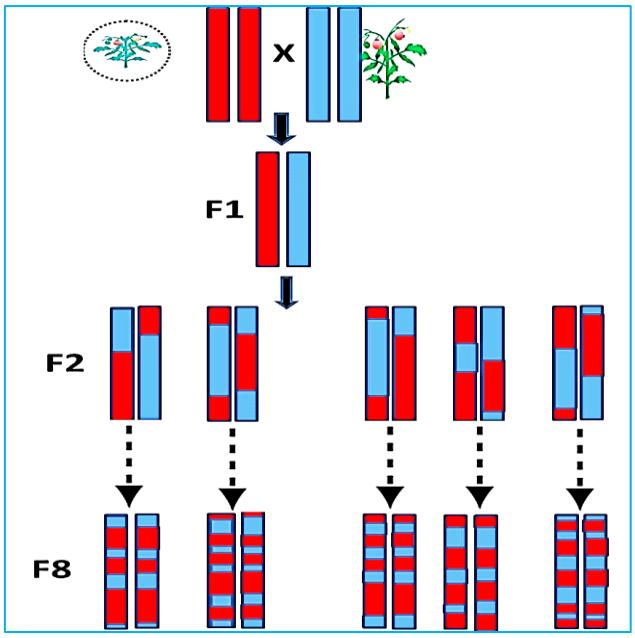

Fig 1 Self-pollination and used single seed descended procedure to develop F_8 RIL population

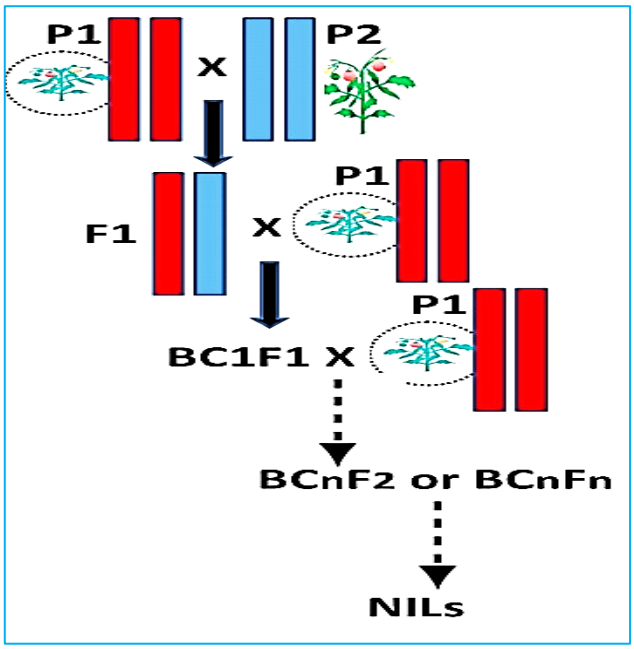

Fig 2 Strategy for developing Near Isogenic Lines (NIL)

Statistical Methods for QTL Mapping

Multiple QTL Mapping (MQM), Composite Interval Mapping (CIM), Single Marker Approach (SMA), Simple Interval Mapping (SIM), and Multi trait Interval Mapping (MIM) approaches are commonly used to analyze QTL and trait association studies [20-24]. The SMA is also referred called as single point analysis or a single factor analysis of variance. It is a one of the methods employed for quickly scanning the entire genome to find out the best QTLs. As part of this SMA analysis each molecular marker locus that are free from other loci and mostly this practice is incapable to identify quantitative trait loci position. The F-test is used to evaluate if there are any substantial differences between genotype groups. The method cannot determine whether the markers are associated with one or more quantitative trait loci; the possibility of quantitative trait loci detection decreases with distance among the molecular marker and quantitative trait loci; the effects of quantitative trait loci are undervalued due to confusing with recombination frequencies; and its accuracy is lower when compared to other methods.

Lander and Botstein [22] proposed Simple Interval Mapping (SIM) methodology and it is based on linkage maps and this approach is known as the two-marker method. QTL is determined in this analysis by generating intervals between two markers at various points. It outperforms the Single Marker Approach (SMA approach but falls short of the Composite Interval Mapping (CIM) and Multi trait Interval Mapping (MIM) methodologies in terms of accuracy. The likelihood ratio test is used in this technique to evaluate every quantitative trait loci in the interval generated by adjacent markers. Simple Interval Mapping (SIM) is commonly used because it is simple to perform using statistical packages such as MAPMAKER/QTL [25]. Lander and Botstein [22] devised equations for significance levels acceptable for interval mapping including the required overall false positive rate, genome size, number of chromosomes, number of marker intervals. However, SIM will not account for genetic variation resulting from other QTLs when different QTLs are segregating in a cross. The limitations of SIM in this situation are the same as those of single marker analysis.

To address some of the limitations with SIM, Composite Interval Mapping (CIM) and Multiple QTL Mapping (MQM) approaches have been developed. The Composite Interval Mapping (CIM) approach was introduced independently by Jansen and Stam [26] and Zeng [23]. MQM is a practical, relevant, and sensitive approach for mapping QTL in experimental populations. MQM is a method of QTL mapping that is associated to Haley-Knott regression [21] and composite interval mapping [23] where it MQM incorporates the advantages of interval mapping with generalized linear model regression [22], [26]. This is accomplished by simultaneously utilizing (part of) the markers as cofactors to reduce the impact of additional QTLs and fitting one QTL at a time in a specified interval. This approach is based on an interval test that combines interval mapping with multiple regression in an effort to extract and separate specific QTL effects. To determine genetic variance due to non-target QTLs, the partial regression coefficient is used. In each QTL analysis, a marker interval and a few other single markers are taken into account. The advantages of CIM are as described in the following: i. mapping of multiple Quantitative characters can also be done in one dimension; ii. While using linked markers as covariates, the assessment isn't really affected by QTL out of region, increasing the accuracy of QTL mapping; and iii. by removing as much genetic variance created by other QTL as possible, the residual variance is lowered, thereby improving of QTL identification. CIM is more effective than SIM, but it is not as widely used in QTL mapping. Multiple interval mapping

(MIM) is another statistical method developed for QTLs mapping. It is used to map multiple QTLs and has the ability to detect QTL×QTL interactions. In order to map numerous putative QTL straight into the framework, it uses multiple marker intervals simultaneously. The MIM approach is based on Cockerham's model for interpreting genetic variables and the maximum likelihood method used for estimating genetic parameters.

CONCLUSION

The recognition and localization of genetic markers influencing quantitative traits is critical in plant breeding. Quantitative Trait Loci (QTL) analysis is useful in determining the potential number of loci, their distribution in the genome, and equality of effects. For the preparation of QTL data for analysis, high density genetic or linkage maps developed using various molecular markers are required. QTL maps are generated using linkage analysis between molecular markers and trait association studies, and they can be used for gene pyramiding, germplasm screening for abiotic and biotic stresses, and so on. In conclusion, QTL analysis is useful, in order to determine the potential number of loci, their distribution throughout the genome, the equality of their effects, and the mode of their action. Due to their high polymorphism, availability, and co-dominance nature of molecular markers they are quite advantageous in identification of number and location of QTLs.

LITERATURE CITED

1. Angaji SA. 2009. QTL mapping: A few key points. *International Journal of Applied Research in Natural Products* 2(2): 1-3.
2. Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica* 142: 169-196.
3. Herve D, Fabre F, Berrios EF, Leroux N, Al Chaarani G, Planchon C, Sarrafi A, Gentzbittel L. 2001. QTL analysis of photosynthesis and water status traits in sunflower (*Helianthus annuus* L.) under greenhouse conditions. *Journal of Experimental Botany* 52(362): 1857-1864.
4. Albertson RC, Streelman JT, Kocher TD. 2003. Directional selection has shaped the oral jaws of Lake Malawi cichlid fishes. *Proceedings of National Academy of Science USA* 100(9): 5252-5257.
5. Dudley JW. 1993. Molecular markers in plant improvement: manipulation of genes affecting quantitative traits. *Crop Science* 33: 660-668.
6. George MLC, Prasanna BM, Rathore RS, Setty TAS, Kasim F, Azrai M, Vasal S, Balla O, Gupta PK. 2002. Molecular markers and QTL analysis in crop plants. *Current Science* 83: 113-114.
7. Collard BCY, Pang ECK, Taylor PWJ. 2003. Selection of wild Cicer accessions for the generation of mapping populations segregating for resistance to ascochyta blight. *Euphytica* 130: 1-9.
8. Kearsey MJ. 1998. The principles of QTL analysis (a minimal mathematics approach). *Jr. Exp. Botany* 49: 1619-1623.
9. Yu JK, Mangor L, Thompson KJ, Edwards MB, Slabaugh, Knapp SJ. 2002. Allelic diversity of simple sequence repeat markers among elite inbred lines in cultivated sunflower. *Genome* 45: 652-660.
10. Balyan HS, Gupta PK, Rustgi S, Bandopadhyay R, Goyal A, Singh A, Kumar A, Kumar N, Sharma S. 2005. Development and use of SSRs of bread wheat for genetic and physical mapping and transferability to the species of *Triticum-Aegilops* complex. *Czech Jr. Genetics and Plant Breeding* 3: 141-144.
11. Gupta PK, Balyan HS, Bandopadhyay R, Kumar N, Sharma S, Kulwal PL, Rustgi S, Singh R, Goyal A, Kumar A. 2005. Development and use of molecular markers for wheat genomics. *Annual Wheat Newsletter* 51: 51-54.
12. Parida S, Kumar KAR, Dalal V, Singh NK, Mohapatra T. 2006. Unigene derived microsatellite markers for the cereal genomes. *Theor. Appl. Genetics* 112: 808-817.
13. Tanksley SD, Young ND, Paterson AH. Bonierbale M. 1989. RFLP mapping in plant breeding: New tools for an old science. *Biotechnology* 7: 257-264.
14. Gupta PK, Varshney RK, Sharma PC, Ramesh B. 1999. Molecular markers and their applications in wheat breeding. *Plant Breeding* 118: 369-390.
15. Beckmann J, Soller M. 1986. Restriction fragment length polymorphisms in plant genetic improvement. *Oxford Surveys of Plant Mol. Biol. Cell Biology* 3: 197-250.
16. Davierwala A, Chowdari K, Kumar S, Reddy A, Ranjekar P, Gupta V. 2000. Use of three different marker systems to estimate genetic diversity of Indian elite rice varieties, *Genetica* 108: 269-284.
17. Huettel B, Winter P, Weising K, Choumane W, Weigand F, Kahl G. 1999. Sequence-tagged microsatellite site markers for chickpea (*Cicer arietinum* L.). *Genome* 42: 210-217.
18. Mohapatra T, Singh KS, Swain S, Sharma R, Singh N. 2003. STMS-based DNA fingerprints of the new plant type wheat lines. *Current Science* 84: 1125-1129.
19. Winter P, Pfaff T, Udupa S, Huttel B, Sharma P, Sahi S, Arreguin-Espinoza R, Weigand F, Muehlbauer FJ, Kahl G. 1999. Characterisation and mapping of sequence tagged microsatellite sites in the chickpea (*Cicer arietinum* L.) genome. *Mol. Gen. Genetics* 262: 90-101.
20. Lander ES, Botstein D. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121: 185-199.
21. Haley C, Knott S. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 69: 315-324.
22. Jansen RC. 1993. Interval mapping of multiple quantitative trait loci. *Genetics* 135(1): 205-211.
23. Zeng, Zhao-Bang. 1994. Precision mapping of quantitative trait loci. *Genetics* 136(4): 1457-1468.
24. Zou W, Zeng ZB. 2008. Statistical methods for mapping multiple QTL. *Int. Jr. of Pl. Genomics* 286561. 10.1155/2008/286561.
25. Lincoln S, Daly M. Lander E. 1992. Mapping genes controlling quantitative traits with MAPMAKER/QTL. Whitehead Institute Technical Report, 2nd Edition, Whitehead Institute, MA.
26. Jansen Ritsert C, Stam P. 1994. High resolution of quantitative traits into multiple loci via interval mapping. *Genetics* 136(4): 1447-1455.