Mapping Quantitative Trait Loci in Plants: Approaches and Applications

Plant breeding is the science and art of the genetic improvement of plants through the process of selection for various agronomically important characters. Plant breeding efforts always include an element of chance, because of numbers of genotypes and environments and consequent phenotypes to be evaluated are limitless. However, considerable breakthroughs have been made towards the development of improved cultivars with increased productivity and quality for their food, feed, fibre or aesthetic value. These are possible due to the continuous and dedicated efforts of plant breeders to select a suitable genotype to suit the environment, so that it can grow and yield well to the ultimate user. Thus, selection of the appropriate cultivar remains as one of the key decisions that a plant breeder must make.

Selection methods in plant breeding depend on the natural reproductive system of the crops concerned. They can, therefore, be classified according to crops that reproduce by self fertilization, cross fertilization and by asexual means. The other important factor influencing the process of selection is the type of gene action involved in the expression of the traits to be selected and the resultant nature of variation in those traits. Considering the nature of variation for various traits, it can be classified as: 1) discrete or discontinuous variation- i.e. showing distinct classes with well defined phenotypic expression and 2) non-discrete or continuous variation- i.e. manifesting indistinguishable or varying degree of expression. Most of the economically important traits in crops species are quantitatively inherited. An understanding of selection theory in plant breeding must be based upon knowledge of inheritance of such traits and the final success of selection cannot be treated without a certain degree of mathematical and statistical handling. This paper reviews Marker based Quantitative Trait Loci (MBQTL) analysis and its applications in crop breeding.

Quantitative variation of characters

The mathematical foundations of a study of quantitative variation, first initiated by Galton's application of mathematics to biological problems (Galton, 1889) were laid much before the rediscovery of Mendel’s work. Galton (1989), based on his correlation analysis, could establish that hereditary transmission is equilinear from the two parents leading to the proposal of Law of Ancestral Heredity. Then, Johannsen (1909), Nilsson-Ehle (1909), East (1910) and Fisher (1918) uncovered the specific relationships between Mendelian and early biometrical approaches and that formed the basis for understanding of quantitative genetic variation.

In particular, Johannsen (1909) stated that both heritable and non-heritable factors were responsible for the variations observed in the seed weight of beans, *Phaseolus vulgaris* and that contributions of these factors could be distinguished in progeny tests provided a major insight into the relationship between the genotype and the phenotype. Then, Nilsson-Ehle (1909), investigating grain colour in wheat, discovered that hereditary factors existed whose actions in the determination of the phenotype were similar, if not identical, and realized that a number of such individual actions, could account for the phenotype observed in the continuous variation of quantitative characters. These attempts to analyse traits with continuous variation remained as foundations to analyse most of the
economically important quantitatively inherited traits. An understanding of selection theory in plant breeding must be based on knowledge of inheritance of such traits. Mather (1941) while proposing simplified assumptions of biometrical models to study quantitative genetic variation used the term 'polygenes' for the genes controlling quantitative traits. These polygenes were said to be many and had similar phenotypic effects that were small, relative to other sources of variation.

Though several simplified assumptions and models are made available to analyze the quantitative genetic variation in crop plants, these are not universally established facts. However, the biometrical models/techniques developed over time remained sole analytical tools for the study of quantitative genetic variation of crops until the recent past. The advent of molecular markers and marker based Quantitative Trait Loci (QTL) analysis made several exciting outputs on this vital field of crop breeding. “But one cannot rule out that early biometrical approaches are no more useful and the new and developing marker based QTL analysis is not having any limitation”.

Galton (1889) laid the foundation for the study of quantitative genetic variation. Like Mendel, who worked on characters with contrasting phenotypes, Galton started his analysis on characters with continuous distribution. Unlike Mendel’s work Galton’s was not noticed. Like the rediscovery of Mendel’s laws, Galton’s work was spearheaded and expanded by Pearson (1900) as a ready-made opposition to Mendel’s proposals. This formed the base for the ever lasting conflicts between Mendelians and Galtonians. The major conflicts established on the inheritance of phenotype by Mendelian and Galtonian schools of thoughts are furnished in Table 1.

<table>
<thead>
<tr>
<th>Mendelian School</th>
<th>Galtonian School</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scorable</td>
<td>Measurable/Countable</td>
</tr>
<tr>
<td>Alternate</td>
<td>Variable</td>
</tr>
<tr>
<td>Single factor controlled</td>
<td>Multiple factors controlled</td>
</tr>
<tr>
<td>Results of Macromutation</td>
<td>Results of Micromutation</td>
</tr>
<tr>
<td>Categorical</td>
<td>Continuous or Normal distribution</td>
</tr>
<tr>
<td>Phenotype = Genotype</td>
<td>Phenotype = Genotype + Environment + Genotype x Environment</td>
</tr>
<tr>
<td>High heritability</td>
<td>Low heritability</td>
</tr>
<tr>
<td>Qualitative</td>
<td>Quantitative</td>
</tr>
<tr>
<td>Rapid responses to artificial selection</td>
<td>Slow response to artificial selection</td>
</tr>
<tr>
<td>Marker aided selection is easy</td>
<td>Marker aided selection is not so easy</td>
</tr>
</tbody>
</table>

The studies of Johannsen (1909) showed that seed weight in *P. vulgaris*, displayed quantitative variation, and that this variation could arise from non-heritable causes as well as heritable factors. His studies further revealed that this variation from both heritable and non-heritable sources could not be separated by mere visualization but only by breeding tests. At the same time, Nilsson-Ehle (1909) showed that the continuous variation observed for the seed colour in wheat could be attributed to the control by many genes
with equal effects. In other words, quantitative variation that Galton observed was due to systems of genes, each inherited in the Mendelian fashion with similar phenotypic effects. With this background information, scientists started developing methods to analyze quantitative traits for the improvement of crop species. A detailed review on understanding quantitative genetic variation is made available by Mackay (2001) and Barton and Keightley (2002).

Locating QTLs using Molecular Marker Linkage Maps

QTLs can be of three types viz., 1) major QTLs, 2) major + minor QTLs and 3) minor QTLs. Usually major QTLs control qualitative traits and have the Mendelian inheritance whereas the other two types deviate the Mendelian nature of inheritance and make the situation difficult to trace them. The different types of QTLs and their interaction with environment to produce a phenotype are illustrated in Fig 1.

![Diagram](image)

Fig 1. Manifestation of a phenotype and the number of genes involved under varying environmental conditions. Source: Xu (1997)

Linkage between a genetic marker and QTL was first demonstrated by Sax (1923) by associating the seed size (a quantitative trait) with seed colour (a morphological marker) in *P. vulgaris*. However, associating specific QTLs with genetic marker(s) and locating them on linkage groups was a great task by the conventional genetic analysis, since the non-availability of more genetic markers remained one of the major practical limitations (Thoday, 1961). This limitation was remedied by the construction of saturated molecular marker maps, permitting systematic searches of an entire genome for QTLs influencing a trait (Paterson et al. 1988). The advent of molecular markers provided the geneticists with powerful new tools for identifying the component Mendelian loci of those complexly inherited traits. The process of QTL analysis requires 1) a suitable mapping population segregating for the trait(s) of interest, 2) a saturated linkage map of molecular markers, 3) a suitable phenotypic screening method to quantify the level of trait’s manifestation and 4) a set of powerful statistical packages to identify the QTLs. Steps involved in mapping QTLs involving a mapping population is shown in Fig. 2.
Suitable mapping population
It would be always advantageous if accurate predictions are made by using early generations (e.g. F2, F3, BC population etc). However, the predictions made involving early generations would be misleading because of camouflaging effect in early generation of the major gene on many other minor genes. This camouflaging effect of major genes on minor genes can be eliminated by continuous inbreeding to evolve recombinant inbred lines (RILs) (Allard and Harding, 1963). Ways to create various kinds of mapping populations suitable for mapping QTLs are shown in Fig. 3

Thus, RILs can remain as the best choice of population for QTL analysis. As an alternative step doubled haploid lines (DHL) can also be used. The inherent homozygosity prevailing in the individuals of these two populations make the RILs and DHLs as immortals and help to have as many replications as required by the experiment.
Saturated linkage map
Thoday (1961) emphasized that the main practical limitation in localizing QTLs, seems to be the availability of suitable markers. This limitation was remedied by the advent of DNA markers - polymorphic and locus characterized by a number of variable lengths paved the way for easy and faster linkage map construction. The construction of saturated linkage maps with different types of DNA markers facilitated the systematic searches of an entire genome for QTLs influencing a particular trait in tomato (Bernatzky and Tanksley, 1986; Paterson et al 1988). Subsequent to this study linkage maps have been constructed with DNA markers in maize (Helentjaris et al. 1986), lettuce (Landry et al. 1987), rice (McCouch et al. 1988), potato (Bonierbale et al. 1988), wheat (Chao et al. 1989; Kam-Morgan et al. 1989) and common bean (Vallejos et al. 1992). Most of these maps were based on RFLP markers (Grodzicker et al. l974). The technological advances in molecular biology led to the evolution of DNA markers (Schlotterer, 2004). These maps were supplemented with more DNA marker types such as randomly amplified polymorphic DNA (Williams et al. 1990; Welsh and McClelland, 1990), inter-simple sequence repeats (Zietkiewicz et al. 1994), amplified fragment length polymorphisms (Vos et al. 1995) and simple sequence repeats (Beckman and Soller, 1990). Several saturated linkage maps of major crop species were constructed integrating various types of DNA markers (Maheswaran et al. 1997, Cho et al. 1998; Harushima et al. 1998, Price et al. 2000; Bhattaramakki et al. 2000; Haussmann et al. 2002; Song et al. 2004, Xia et al. 2007). Presently SSR markers are being considered as the markers of choice for linkage map construction in almost all the crops: sorghum (Taramino et al. 1997); wheat (Roader et al. 1998), rice (Temnykh et al. 2000) and maize (Sharopova et al. 2002). These path breaking attempts paved the way for the construction of several linkage maps of molecular markers in all most all the crop species exclusively for locating the QTLs of various agronomic traits.

Reliable phenotypic screening and generation of phenotypic data
Phenotype of a trait is always under the dynamic change. Measuring the phenotype may not give same measures when the trait is under the control of several genes that too genes having positive and negative controls with known/unknown interactions. To adequately explore the QTL during the mapping phase, the phenotype must be evaluated in replicated trials in different environments. Moreover, phenotypic screening should be done based on reliable and reproducible screening methods. Large data sets can be generated by the coordinated efforts of several groups, providing valuable information about genes governing quantitative characters in a range of environments. Such data will provide information about the magnitude of the effect of different QTL and whether there is interaction between QTL and environment (McCouch, 1993).

The issues related to population development and construction of linkage maps does not pose many problems with the existing level of knowledge. Though, issues related to methods for detecting QTL, software for QTL analysis are having problems such as inaccurate detection of QTL (occurrence of false positives and false negatives), the issues associated with phenotypic screening pose severe threat to an emerging tool of plant breeding.

Most agronomically important characters involve multiple genes that interact with each other and with the environment in complex ways. This creates a situation wherein QTLs can be detected only some of the time. This necessitates designing of experiments to
qualify the value of specific QTL. To adequately explore the value of QTLs, the phenotypes must be evaluated in well-replicated trials in different environments. The conduct of replicated trials warrants a suitable population which can be effectively replicated. Here comes the problem of developing an immortal population such as RILs and DHLs. Developing both kinds of populations and their maintenance is a cumbersome process even by a well established breeding institute. QTLs are hypothetical genes based on statistical inference. Genetic effects used in QTL mapping could have very little biological meaning. To have a biological meaning of QTL mapping the selection of traits to be phenotyped is very important. Moreover, phenotypic manifestations of several traits are not amenable for QTL detection due to the non-availability of methods for proper quantification. Another important issue is about the logic behind the selection of traits for QTL detection. The detection of QTL is a trait dependent process. In other words, all the traits are not measurable and all possible measurable traits need not variable.

Mapping methods and software

*Identification and characterization of QTLs*

Identification of QTLs for a trait has several steps. The first one is genotyping the individuals of a population by molecular marker survey. One can get three possible genotypes for each marker. *e.g.* (A/A, A/a and a/a). The second step is phenotyping the population for the trait of interest. Grouping the individuals based on the genotype of each marker and finding out the group mean is the third step. Fourth step involves carrying out ANOVA to determine whether differences between the individual groups of each marker are significant. If there is no significance, it indicates the absence of QTL near the marker. The significance indicates presence of a QTL associated with the marker. The identified QTLs can be of any of the following types and it depends upon the nature of the trait (Fig. 4)

![Identification of QTL based on significant difference between groups of phenotypes associated with specific marker genotypes](image)

Fig 4. Identification of QTL based on significant difference between groups of phenotypes associated with specific marker genotypes
QTL mapping methods
Conceptually, QTL mapping amounts to a three-step recipe: scan the entire genome with a dense collection of genetic markers; calculate an appropriate linkage statistic $S(x)$ at each position $x$ along the genome; and identify the regions in which the statistic $S$ shows a significant deviation from what would be expected under independent assortment. The underlying assumptions of QTL mapping involving molecular markers are: 1) genes controlling quantitative traits are located on the genome, just like simple genetic markers, 2) if the markers cover a large portion of the genome then there is a large chance that some of the genes controlling the quantitative traits are linked with some of the genetic markers. 3) if the genes and markers are segregating in a genetically defined population, then the linkage relationship among them may be resolved by studying the association between trait variation and marker segregation pattern. The association between quantitative trait variation and marker segregation pattern can be carried out by the following methods. The basis of all QTL detection, regardless of the crop, is the identification of association between genetically determined phenotypes and specific genetic markers. Mapping QTLs using molecular linkage maps was well explained by Knapp et al. (1990). The possible methods of analysis to detect QTL include: 1) single marker analysis (otherwise called as Marker-Trait (MT) Method) and 2) interval analysis.

Single marker analysis
The single marker analysis (SMA) is a good start not only for learning QTL mapping, but also for practical data analysis. Single marker analysis is the method used in earliest studies on QTL mapping (Edwards et al., 1987; Weller et al., 1988). In this method, one marker is involved at a time to find the QTL-marker association. The single marker analysis can be implemented as a simple t-test, ANOVA, linear regression, and likelihood ratio test and maximum likelihood estimation (Haley and Knott, 1992; Nienhuis et al., 1987; Wang et al., 1994). Single marker analysis is simple in terms of data analysis and implementation. It can be performed using common statistical software. Marker orders and linkage map with adequate number of markers are not required.

![Fig. 4: Association of a marker with a putative QTL](Source: Liu, 1998)

The disadvantages of the single marker analysis are: 1) the putative QTL genotypic means and QTL positions are confounded. This confounding causes the estimated QTL effects to be biased and the statistical power to be low particularly when linkage map density is low and 2) QTL positions cannot be precisely determined, due to the non-dependence among the hypothesis tests for linked markers that confound QTL effect and position.
Interval analysis or Interval mapping

Interval mapping (IM) is considered as a second level of QTL mapping. QTL mapping by this method requires prior construction of a marker based linkage map. The interval mapping approach is based on the joint frequencies of a pair of adjacent markers and a putative QTL flanked by the two markers. Interval mapping methods for QTL detection evolved over the period into 1) simple interval mapping (SIM), 2) composite interval mapping (CIM) and 3) multiple interval mapping (MIM).

![Fig 5. Association of a putative QTL to two flanking markers (Source: Liu, 1998)](image)

Simple Interval Mapping (SIM)

Simple interval mapping (SIM) considers one QTL at a time in the model for QTL mapping. Therefore, SIM can bias identification and estimation of QTL when multiple QTL are located in the same linkage group (Zeng, 1994). SIM evaluates the association between the trait values and the expected contribution of hypothetical QTL (target QTL) at multiple analysis points between each pair of adjacent marker loci (the target interval). The expected QTL genotype is estimated from the genotypes of flanking marker loci and their distance from the QTL. Since there is usually uncertainty in the QTL genotype, the likelihood is sum of terms, one for each possible QTL genotype, weighted by the probability of that genotype given the genotypes of the flanking markers. The analysis point that yields the most significant association may be taken as the location of a putative QTL. Although IM represented a significant contribution to QTL analyses, it is based on the null hypothesis of no QTL: an incorrect assumption for quantitative traits.

Composite Interval mapping (CIM)

Both SMA and IM are biased when multiple QTL are linked to the marker/interval being considered. To deal with multiple QTL problems, Jansen (1993) Rodolphe and Lefort (1993) and Zeng (1993) independently proposed the idea of combining SIM with multiple regression analysis in mapping. Multiple regression methods were integrated with IM to increase the probability of including all significant QTLs in the model. This method was named as composite interval mapping (CIM).

CIM evaluates the possibility of a target QTL at multiple analysis points across each intermarker interval. However, at each point it also includes the effect of one or more background markers, as defined in SIM. The inclusion of a background marker in the analysis helps in one of two ways, depending on whether the background marker and the target interval are linked. If they are not linked, inclusion of the background marker makes the analysis more sensitive to the presence of a QTL in the target interval. If they
are linked, inclusion of the background marker may help to separate target QTL from other linked QTL on the far side of the background marker (Zeng, 1993, 1994).

Multiple Interval Mapping (MIM)
Though CIM produced more accurate and precise estimates than IM, the inclusion of too many cofactors reduced the power to identify QTL relative to IM (Jansen, 1994; Zeng, 1994; Utz and Melchinger, 1994). Kao et al. (1999) proposed new method viz, multiple interval mapping (MIM) to deal with the mapping of multiple QTLs. When compared to SIM and CIM, MIM tends to be more powerful and precise in detecting QTLs MIM method uses multiple marker intervals simultaneously to fit multiple putative QTLs directly in the model for mapping QTLs. The MIM method is based on Cockerham’s model for interpreting genetic parameters and the method of maximum likelihood for estimating genetic parameters. With the MIM approach, the precision and power of QTL mapping could be improved. Also, epistasis between QTLs, genotypic values of individuals and heritability of quantitative traits can be readily estimated and analyzed.

Software for QTL mapping
Compared to general statistical analysis of biological data, statistical analysis for the study of genes controlling complex traits has the following characteristics: 1) many repeated analysis in one task, 2) lack of standard distribution for some test statistics and 3) complexity of models used in QTL mapping. For using these software packages a known linkage map is needed for either running the programmes or interpreting results. Several companion packages are also needed for linkage map construction. These packages are often endowed with problems such as 1) not having user friendly interface, 2) user support is limited due to their non-commercial status, 3) irrelevant statistical or genetic models for the analysis, 4) taking much time for giving the output and 5) inconsistency in results obtained from different software for the same data set. Available public domain software packages for mapping QTLs associated with complex traits are not adequate for establishing conclusive decisions on complex traits in terms of user interface, flexibility and user support. Software packages with commercial quality are needed to accommodate the growing needs of data analysis and management in QTL mapping. The details on some of the software routinely employed in QTL mapping are given in Table 2.
<table>
<thead>
<tr>
<th>Name</th>
<th>Platform</th>
<th>Remarks</th>
<th>References</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mapmaker/QTL Version 1.1</td>
<td>Unix or DOS</td>
<td>Widely used programme; Interval mapping and composite interval mapping; requires Mapmaker/Exp for map construction; freely downloadable</td>
<td>Lincoln <em>et al.</em> 1992a, 1992b</td>
<td>ftp://genome.wi.mit.edu/pub/mapmaker3</td>
</tr>
<tr>
<td>QTL Cartographer V2.5</td>
<td>Unix; DOS; Mac; Windows</td>
<td>Offers several varieties of composite interval mapping; freely downloadable</td>
<td>Basten <em>et al.</em> 1997</td>
<td><a href="http://statgen.ncsu.edu/qtlcart/index.php">http://statgen.ncsu.edu/qtlcart/index.php</a></td>
</tr>
<tr>
<td>Map Manager QTX</td>
<td>Mac</td>
<td>Graphical user interface for data entry and display; Designed as mapping programme and data preparation programme for other mapping programmes</td>
<td>Manly <em>et al.</em> 2001</td>
<td><a href="http://www.mapmanager.org/">http://www.mapmanager.org/</a></td>
</tr>
<tr>
<td>QGene 4.0</td>
<td>Mac</td>
<td>Graphical interface for displaying trait data; marker-trait relationship and marker genotype</td>
<td>Nelson, 1997</td>
<td><a href="http://coding.plantpath.ksu.edu/qgene">http://coding.plantpath.ksu.edu/qgene</a></td>
</tr>
<tr>
<td>PLABQTL</td>
<td>DOS; AIX</td>
<td>QTL detection and environmental effect</td>
<td>Utz and Melchinger, 1996</td>
<td><a href="https://www.uni-hohenheim.de/plantbreeding/software/">https://www.uni-hohenheim.de/plantbreeding/software/</a></td>
</tr>
<tr>
<td>Software</td>
<td>Platform</td>
<td>Description</td>
<td>Reference</td>
<td>Website</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------------------</td>
<td>----------------------------------------------</td>
</tr>
<tr>
<td>MQTL v 0.98</td>
<td>DOS; SUN</td>
<td>Simplified Composite Interval Mapping Estimation QTL x Environment effect</td>
<td>Tinker and Mather, 1995a, 1996b</td>
<td>ftp://gnome.agrenv.megil.ca/pub/genetics/software/MQTL/</td>
</tr>
<tr>
<td>Multimapper</td>
<td>Unix</td>
<td>QTL mapping based on single linkage group; Provision for comparing threshold statistic generated by other programmes for QTL mapping</td>
<td>Sillanpaa, 1998</td>
<td><a href="http://www.RNI.Helsinki.FI/~mjs/">http://www.RNI.Helsinki.FI/~mjs/</a></td>
</tr>
<tr>
<td>MultiQTL</td>
<td>Windows XP</td>
<td>Easy data input; Fast bootstrapping and permutation test</td>
<td>MultiQTL LTD</td>
<td><a href="http://www.multiqtl.com/">http://www.multiqtl.com/</a></td>
</tr>
<tr>
<td>Software</td>
<td>Operating System</td>
<td>Description</td>
<td>Reference</td>
<td>Website</td>
</tr>
<tr>
<td>------------</td>
<td>------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td>-----------</td>
<td>----------------------------------------------</td>
</tr>
<tr>
<td>QTL Express</td>
<td>Windows XP</td>
<td>Tools for permutation analysis to set significance levels and bootstrap analysis to estimate confidence regions for the QTL location are provided</td>
<td>Seaton et al. 2002</td>
<td><a href="http://qtl.cap.ed.ac.uk/">http://qtl.cap.ed.ac.uk/</a></td>
</tr>
<tr>
<td>Grid QTL</td>
<td>Windows XP</td>
<td></td>
<td>Seaton et al. 2006</td>
<td><a href="http://www.gridqtl.org.uk/">http://www.gridqtl.org.uk/</a></td>
</tr>
<tr>
<td>The QTL Cafe</td>
<td>Multiple OS</td>
<td>The QTL café is a JAVA based package which provides a user friendly way to perform QTL analyses</td>
<td></td>
<td><a href="http://www.biosciences.bham.ac.uk/labs/kearsey/">http://www.biosciences.bham.ac.uk/labs/kearsey/</a></td>
</tr>
<tr>
<td>QTLBIM</td>
<td>Windows XP</td>
<td>Provides a Bayesian model selection approach to map multiple interacting QTL</td>
<td></td>
<td><a href="http://www.qtlbim.org/">http://www.qtlbim.org/</a></td>
</tr>
<tr>
<td>R/qtl</td>
<td>Windows XP</td>
<td></td>
<td></td>
<td><a href="http://www.rqtl.org/">http://www.rqtl.org/</a></td>
</tr>
<tr>
<td>MetaQTL</td>
<td>Windows XP</td>
<td>Java package designed to perform the integration of data from the field of gene mapping experiments</td>
<td></td>
<td><a href="http://www.bioinformatics.org/mqt/">http://www.bioinformatics.org/mqt/</a> wiki/M ain/HomePage</td>
</tr>
<tr>
<td>MapQTL 6</td>
<td>Windows XP</td>
<td>Analyzing QTL experiments with interval mapping, with the powerful MQM mapping</td>
<td></td>
<td><a href="http://www.kyazma.nl/">http://www.kyazma.nl/</a></td>
</tr>
<tr>
<td>J/qtl</td>
<td>Windows XP</td>
<td></td>
<td>The Jackson Laboratory</td>
<td><a href="http://research.jax.org/faculty/churchill/research/qtl/index.html">http://research.jax.org/faculty/churchill/research/qtl/index.html</a></td>
</tr>
</tbody>
</table>
Power, precision and accuracy of QTL mapping

QTL analysis includes three stages: 1) detection, 2) mapping and 3) fine mapping. Detection and mapping (assigning chromosomal location) are often accomplished simultaneously, but they are logically and statistically distinct (Beavis, 1998).

**Power of detection:** Power is the probability of identifying a QTL of known magnitude, given the predetermined frequency of false positive association ($\alpha$). Each QTL detection experiment provides an estimate of the strength of a QTL. In some experiments, the QTL will be over estimated, in others, underestimated. This variability may determine whether the QTL appears to be statistically significant, that is, whether the QTL is detected in that experiment or not. The power of a QTL detection experiment, at a given level of statistical significance depends upon the strength of the QTL and the number of progeny in the population.

The strength of the QTL can be determined based on the fraction of the total trait variance that it explains. Those, which explain over 20 percent of the variance, are strong QTL; traits controlled by such QTL can be considered almost Mendelian. At the other extreme, weak QTL, which explain 1 percent or less of the trait variance, require at least a thousand progeny to detect them with high power. Detection of such QTL is routinely feasible. Between those extremes are moderate QTL, which can be detected with crosses of reasonable size but not necessarily at high power.

**Precision of mapping**

Precision is a measure of the dispersion of repeated independent estimates of genomic positions or genetic effects of the alleles at a QTL and reported by inverse measures such as standard errors or confidence intervals. The size of a confidence interval is expected to be inversely proportional to the number of progeny in the mapping population and to the square of the strength of the QTL. Weak QTLs, as defined above, can be assigned to a chromosome, but not located with more precision. Strong QTLs can be located by a large backcross or intercross in confidence intervals as small as 1 cM. For strong QTLs, precision is limited by the number of recombinants in backcross or intercrosses and QTL can be located more precisely in advanced intercrosses (Darvasi, 1998). The power and precision of QTL mapping depends on the test statistic derived based on the asymptotic distribution. When single marker analysis is involved in QTL mapping, ‘t’ and ‘F’ statistics are used to assess the power and precision of contrasting marker genotypic classes (Soller, 1976; McMillan and Robertson, 1974).

**Accuracy of mapping**

Accuracy is a measure of how close the estimates are to the true values. In practice, accuracy is very difficult to estimate for experimental results because the true values are unknown.

**Test statistic or threshold statistic for claiming QTL detection**

The QTL, by definition, are merely significant statistical associations. These significant associations are detected by having suitable test statistic, otherwise called as ‘critical value’ or ‘threshold statistic’. The reliability and efficiency of the QTL mapping depend considerably on the validity and relevance of the statistical tests used to detect the presence of QTL. Clear statistical guidelines for the interpretation of linkage results are
needed to avoid a flood of false positive (detection of a QTL when actually it is not present) claims. At the same time, an overly cautious approach runs the risk of causing true hints of linkage to be missed (false negative).

‘Critical value’ or ‘threshold statistic’ is a limit fixed to eliminate the detection of spurious QTL and QTL with smaller effects. Fixing a suitable threshold statistic for each population size will help in improving the power of QTL mapping. In using single marker analysis, test of significance is used as threshold statistic. When analysis of variance is used as method to detect QTL, ‘F’ value is used as threshold statistic. In the same manner, for interval mapping LOD score is used as a threshold statistic. The LOD score summarizes the strength of evidence in favour of the existence of QTL with an effect at a position; if the LOD score exceeds a predetermined threshold (usually LOD score of 3.0 is fixed), the presence of a QTL is inferred (Fig.6). For estimating the LOD score, one has to have the odds ratio, which is the ratio between chance of QTL at a given site and chance of no QTL at a given site.

Threshold statistic adopted for different methods may not have the same strength, resulting in differences in detecting the QTL. Under the circumstance, the threshold statistic of each method has to be evolved to eliminate the discrepancies between methods. Churchill and Doerge (1994) evolved a method to relate the LOD score and F statistic of ANOVA.

When using any statistic of any method, as a criterion in model selection for QTL detection, it is very important to determine the appropriate critical value or threshold value for claiming QTL detection such that correct statistical interference about QTL parameters can be made.

The critical values for claiming QTL detection various corrections are followed. Lander and Botstein (1989) suggested to use the Bonferroni correction and Orenstein-uhlenback diffusion when the linkage maps are sparse dense respectively to determine the critical value. The adjustment of critical values to detect QTLs depends on the number and size
of interval, different levels of heritability, different number of multiple linked or unlinked QTL and linked QTL in the same or opposite direction (Lander and Botstein, 1989; Jansen, 1993; Zeng, 1994). Visscher and Haley (1996) suggested that the critical value should be reduced after a QTL of large effect has been detected. and Doerge and Churchill (1996) suggested using permutation test for determining an appropriate critical value for specific data sets. Doerge and Churchill (1996) recommended at least 1000 permutations for establishing a threshold for \( \alpha = 0.05 \). Permutation tests, therefore, can be time consuming and may be impractical on some computers. In the MIM developed by Kao et al. (1999), they were not adopting an appropriate critical value for claiming QTL detection since the significance level of the multiple-QTL model depends on the marker data structure and the unknown true model.

Fixing a correct critical value to detect QTL is still a debatable issue in QTL mapping. Having a uniform stringent standard such as a critical value based on a whole genome search or a critical value based on an infinitely dense map is not acceptable since some times QTL mapping involves few markers (or few chromosomal regions) or a sparse map. For this Lander and Kruglyak (1999) suggested to have a hierarchical search – in which one performs a genome scan with a sparse map and then follows up interesting regions with a denser map. They also proposed the following classification based on the number of times that one would expect to see a result at random in a dense, complete genome scan:

- **Suggestive linkage**: statistical evidence that would be expected to occur one time at random in a genome scan.
- **Significant linkage**: statistical evidence expected to occur 0.05 times in a genome scan (that is, with probability 5 percent).
- **Highly significant linkage**: statistical evidence expected to occur 0.001 time in a genome scan.
- **Confirmed linkage**: significant linkage from one or a combination of initial studies that has subsequently been confirmed in a further sample, preferably by an independent group of investigators. For confirmation, a nominal \( p \) value of 0.01 should be required.

**Trait based QTL mapping**

QTL mapping is a time consuming process considering the larger populations for which molecular marker analysis has to be made. Though the molecular marker analysis is made simple, one cannot ignore the cost involved in molecular marker analysis. This warranted to develop an alternative strategy to map the QTLs of a trait in less time and at low cost and is called trait based QTL mapping. Trait based QTL mapping involves selective genotyping of distributional extremes for trait in a larger segregating population using the molecular markers and then to determine the marker-QTL linkage (Lebowitz et al., 1987). This strategy saves the time and resources in assaying molecular markers. Selective genotyping of 74 lines from a recombinant inbred population of rice involving IR74 and FR13A by using the phenotypic extremes of for submergence tolerance in rice. A linkage map with 202 AFLP markers was constructed and a QTL contributing towards the submergence tolerance was mapped to chromosome 9 of rice (Nandi et al., 1997). The trait based QTL mapping is effective only for one trait at a time because different individuals will likely be in the extremes for different traits. Likewise, it is likely to be effective only for traits controlled by a small number of QTL which show little interaction.
Applications of QTL mapping

The development of saturated linkage maps with DNA markers facilitated to locate the genes of both simple and complex traits. The role of DNA markers in crop improvement, especially to understand the genetic architecture of complex plant traits is note worthy. Paterson et al. (1991) made detailed review on this aspect. Subsequent to this review, Tanksley (1993) presented the strategies for detecting polygenes and their characterization. Even with all these information, there are several issues which are to be addressed in evaluating the usefulness of QTLs in practical plant breeding. QTL analysis is essentially a “black box” approach to dissect complex phenotypes (McCouch, 1993). The approach allows us to identify individual loci showing a statistically significant association with phenotypes of interest, but offers no information about what genes at QTLs code for and what specific functions they serve. QTL studies suggest chromosomal regions where genes controlling quantitative characters are believed to reside, but a single QTL analysis does not present conclusive evidence that this is so. Since, QTL analysis is considered as the method of choice to study complex traits, refinements are made to set right the problems associated with QTL analysis, so that the complex process of crop improvement can be made simpler to a great extent. Though, there are practical constraints in QTL mapping, researchers started exploring the possibilities of applying the QTL mapping in practical plant breeding and related areas. The three major areas where the knowledge of QTL mapping can be employed include: 1) genetic analysis, 2) marker assisted selection (MAS) and 3) introgression breeding.

Genetic analysis

Marker based QTL analysis remains as an important tool in modern day plant breeding to find out the number of genes responsible for the expression of trait and locate them on specific linkage groups. It also helps to detect the nature of gene action, linkage or pleiotropy, epistasis, transgressants and their genetic basis, genetic basis of heterosis and genotype x environment interaction.

Gene action

The phenotypic effect of a gene depends upon the number and nature of alleles present in an organism. The nature of gene action is usually studied involving a hybrid with its parents having two different alleles with distinct phenotypes. Among them, the hybrid is heterozygous carrying one copy of each parental allele. The following table explains possible types of gene action.

<table>
<thead>
<tr>
<th>Phenotype of the hybrid</th>
<th>Relationship between alleles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate between two parents</td>
<td>Additive ($d = 0$)</td>
</tr>
<tr>
<td>Identical to one of the parents</td>
<td>Dominant ($d = a$) or recessive ($d = -a$)</td>
</tr>
<tr>
<td>Incremental effect over both the parents</td>
<td>Overdominant ($a = 0, d &gt; 0$)</td>
</tr>
</tbody>
</table>

In the conventional genetic analysis finding alternative alleles for each trait is a great task. The molecular marker alleles can be associated to each of the QTLs responsible for a phenotype. Based on this association, the average phenotype of individuals carrying ‘no
copy’ ‘one copy’ or ‘two copies’ of the marker allele can be estimated. The change in phenotype per additional copy of the allele is whether towards additivity, dominance/recessive or over-dominance can be found out by adopting specific statistical analysis. Some QTLs cannot be clearly assigned to any of these classes, as they appear intermediate between two classes (partial dominance; \( a > 0, d > 0 \)) (Paterson et al., 1991). For determining the gene action, the ideal population is an \( F_2 \) population having the genotypes segregating in a Mendelian fashion. The back cross populations or recombinant inbred populations are not suitable for determining the gene action, since these populations will not have all the three Mendelian genotypes (no recessive genotypes in back cross population; no heterozygotes in recombinant inbred population). Molecular marker facilitated investigation to understand the types of gene action in maize was proposed by Edwrads et al., (1987) using two \( F_2 \) populations viz., CO159/Tx303 and T232/CM37. About 82 quantitative traits were measured for each plant throughout the season. Approximately 50% of the cases in each population exhibited additive or partial dominant in expression. Another 25% exhibited partial dominance or dominance and the remaining 25% showed over-dominance. Using the same populations, another study was conducted to find out the factors influencing the yield and its components. For plant grain yield, top ear grain weight and ear length, the gene action was primarily dominant or over-dominant (Stuber et al., 1987). Like this types of gene actions for various quantitative traits have been predicted using the Marker based QTL analysis. Results from these studies should prove to be useful for manipulating QTLs in marker facilitated selection programmes.

Linkage or pleiotropy
The phenotypic expression of a trait can be due to the effect of a single gene or sometimes due to the combined effect of closely linked genes. Occasionally a single locus may have control over different traits and the phenomenon is known as pleiotropism. Distinguishing linkage from pleiotropy is not possible in many occasions and the marker based QTL mapping facilitates to determine whether an association between different traits is due to several linked QTLs or to a single QTL with inseparable effects. Paterson et al., (1988) established genomic regions exerting effects on more than one trait (chromosome 6 for fruit weight, soluble solid concentration and pH) in tomato and at least some of the effects are due to pleiotropy. However, they could not distinguish effects of genomic region is due to linkage of QTLs or pleiotropy because of insufficient data. Weller et al., (1988) discovered several of the marker associated main effects in tomato were due to linkage rather than pleiotropy. The QTL mapping study in maize by Edwrads et al., (1992) allowed detecting linkage between QTLs affecting the same traits and separate regions previously found to affect a number of traits (pleiotropy) as QTLs affecting different traits linked together.

Epistasis
Epistasis is a phenomenon of interaction between many different genes and is common in the manifestation of many of the complex traits. Understanding the role of epistasis in the expression of a particular trait will help the breeder to subject the population for further manipulation. Lesser the epistasis on the overall phenotype, easier is the manipulation for that trait. Edwrads et al., (1987) established significant digenic epistasis between pairs of marker-linked genomic regions linked to 15 traits out 82 traits analysed. The results QTL mapping for resiatnce to blast in rice using a recombinant inbred population identified
epistatic effects among loci associated with partial resistance to blast fungus. The two-way interaction analysis in all pair-wise combinations between putative QTLs revealed ten significant epistatic interactions (Wang et al., 1994). Using the same population, Maheswaran (2000) established interactions between flowering QTLs and identified QTL responsible for early flowering, late flowering and flowering per se. Li et al., (1997) reported epistasis for three grain yield components (1000 kernel weight, grain number per panicle and grain weight per panicle) in Lemont/Teqing rice cross.

Transgressive segregation
Transgressive segregation is defined as the appearance of individuals in a segregating population that fall beyond their parental phenotypes (Tanksley, 1993). It may be due to any of the following: de novo mutation, complementary action of genes from the parental lines and unmasking of recessive genes (Rick and Smith, 1953). Transgression is observed in cases where the parents are similar in their means. Transgressive individuals may possess characteristics that will allow them to new ecological nitches (Lewontin and Birch, 1966). de Vicente and Tanksley (1993) carried out QTL mapping for 11 traits in an interspecific cross of taomato. Out of 11 traits, there was evidence for transgressive segregation in eight traits. The QTL mapping using the molecular markers established the fact that the occurrence of significant transgression for these traits was due to the presence of complementary QTL alleles from the two parental lines.

Heterosis
Hybrid vigour (Shull, 1908; East, 1908) or Heterosis (Shull, 1914) is a phenomenon related to heterozygosity resulted by mating two individuals. The level of heterosis is usually low when the parents are closely related. There are two theories on the genetic basis of heterosis: dominance theory (Bruce, 1910, Keeble and Pellew, 1910) and over dominance theory (Shull, 1908; East, 1908). It remains difficult to distinguish between dominance and over dominance as a basis of heterosis. However, molecular markers are expected to throw light on the understanding of genetic basis of heterosis. Identification of genetic factors contributing to heterosis in a hybrid using molecular markers was made in maize (Stuber et al., 1992). The experimental materials were developed by intercrossing two maize inbred lines viz., B73 and Mo17. A total of 264 F2 plants were advanced to F3 generation and a single F3 plant was in turn advanced to F4. Genotyping of F3 was done by bulking 10 F4 plants from each of the F3 plants advanced to F4. Individual F3 plants were then backcrossed to each of the two parental lines to produce progeny which were phenotyped in field evaluations. QTL mapping for grain yield revealed six QTLs in backcross to B73 and eight QTLs in backcross to Mo17. Whenever a QTL for grain yield was detected, the heterozygote had a higher phenotype than the respective homozygote (with only one exception) suggesting not only overdominance but also these detected QTLs played a significant role in heterosis and no convincing evidence for epistasis was found. Based on this study, Xiao et al., (1995) attempted to identify QTLs contributing towards hybrid vigour in an indica/Japonica cross (9024/LH 422) of rice. A total of 37 QTL were identified for 12 quantitative traits. Out 37 QTLs, 27 were detected on only one backcross population. In most of these cases, the heterozygotes were superior to the respective homozygotes. The remaining 10 QTLs were detected in both the backcross populations and the heterozygote had a phenotypes falling between those of the two homozygotes and in no instances were heterozygotes found to be superior to both homozygotes. The results from this study revealed that dominance complementation is
the major genetic basis of heterosis in rice. Yu et al., (1997) reported that the genetic basis of heterosis in rice as a phenomenon due to epistasis. A total of 32 QTLs were found for the four traits. Over-dominance was observed for most of the QTLs for yield. Correlations between marker heterozygosity and trait expression were low, indicating that overall heterozygosity made little contribution to heterosis. Thus, QTL mapping using molecular markers provide an opportunity to unravel many new things about the phenomenon of heterosis and the preliminary studies indicate the heterosis is not only a phenomenon due to dominance or over-dominance, and also due to epistasis. In other words, the genetic basis of heterosis may be varying in different crosses of the same crop and from crop to crop.

Genotype x Environment interaction

Genotype x Environment interaction (G x E interaction) is an essential issue in the assessment of mechanisms of inheritance as well as the prediction of performance in breeding programmes because genotypic values must be inferred from phenotypic responses. The process of QTL mapping offers an opportunity to detect the QTLs for a trait and in turn identify the QTLs which are stable across environments by testing an immortal mapping population (recombinant inbred population or dihaploid population). First attempt to study the QTL mapping across environments was made by Paterson et al., (1991) in an interspecific cross of tomato involving Lycopersicon esculentum and L. cheesmanii. A total of 350 F\textsubscript{2} individuals (location I) and their F\textsubscript{3} progenies (location II & III) were phenotyped for fruit size, soluble solids concentration and pH in three different environments. A total of 29 QTLs were identified in the three environments. Among these 29 QTLs, 4 were detected in the three environments, 10 in two environments and 15 only in a single environment. From this study, it is clear that individual QTLs appear to show a range of sensitivities to environments. In general, QTLs which function consistently over a range of environments are preferred for breeding and the QTL mapping provides detection of stable QTLs across environments to use the information in practical plant breeding. Like this, several QTL mapping studies across environments have been carried out in various crops (Champoux et al., 1995, Maheswran et al. 2000; Stuber et al., 1992; Yu et al., 1997).

Marker Aided Selection

In plant breeding, selecting the best genotype from the variety of genotypes remains the primary objective. Various selection methods were employed by the breeders involving biometrical, mutational and cytogenetical approaches. To the present day breeders, molecular markers are going to be the best tool of selection opening up ways for a separate field of study, the molecular breeding.

In MAS, the tight linkage of marker to a gene is exploited for indirect selection of traits in a breeding programme. Two prerequisites for adopting MAS in breeding programmes are: 1) a tightly linked marker to the gene of interest and 2) a population which is polymorphic for the marker and the gene which are in extreme linkage disequilibrium. Several aspects regarding MAS have been discussed by Stuber (1989) and Melchinger (1990). In plant breeding, two distinct methods are followed: 1) for germplasm improvement (recurrent selection) and 2) for cultivar or hybrid development. These two applications are separated because recurrent selection usually is applied to random mating populations possibly at or near linkage equilibrium, whereas cultivar or hybrid
development typically begins with populations derived by crossing elite inbred lines or near maximum linkage disequilibrium. Lande and Thompson (1990) and Lande (1992) investigated the efficiency of MAS for both individual and mass selection in random mating population.

There are three possible approaches to applying MAS in plant breeding: 1) selection based on markers alone with no measurement of phenotype, 2) simultaneous selection on markers and phenotype and 3) two stage selection, the first stage involving use of markers to select among the genotypes and second involving phenotypic selection among the selected genotypes. The potential efficiency of MAS depends on the heritability of the trait, proportion of genetic variance explained by the markers and the selection method. MAS has been well demonstrated for many of the major gene controlled traits using molecular markers. However, the utility of MAS in manipulating the quantitative traits is still in its early stage and the attempts are sporadic. Dudley (1993) presents the potentials of molecular markers in the manipulation of genes affecting quantitative traits in a detailed manner.

In plant breeding, markers can be applied for the identification of QTLs for trait that is under the control of small number of major genes with large environmental variance, a large number of minor genes with small effects or a mixture of the two. As already explained, for identifying the QTLs of a trait, there should linkage disequilibrium between marker and QTL and for exercising MAS there should not be crossing over between the marker allele and the QTL allele. MAS, when only small number of QTLs is involved, are similar to selecting for qualitative traits. Tanksley et al., (1981) suggested use of isozymes to select for the genotype of a recurrent parent in a backcrossing. Paterson et al., (1988) suggested the use of marker information could reduce the number of backcrosses required by half. During the process of MAS, instead of using a single marker for selection if flanking markers are used, the accuracy of selection improves. When large number of QTLs are to be selected based on markers, then the process of MAS becomes complex. Various approaches of MAS are suggested by Dudley (1993).

Stuber (1989) reported the exploitation of MAS in quantitative trait manipulation. He demonstrated the effectiveness of marker based techniques for identifying and locating QTLs and for detailed genetic investigation of quantitative trait variation. He recorded more precise mapping of QTLs in several plant populations and multiple trait associations with specific genomic regions and demonstrated the transfer of QTLs using MAS for improving the yield level of maize inbred lines.

A total 141 “Enhanced B73” lines are crossed to normal Mo17. In replicated field trials, 45 (32%) of those crosses yielded more than the check hybrid (normal B73 x normal Mo17) by at least 1 SD. Evaluation of crosses of 116 “Enhanced Mo17” lines to normal B73 showed that 51 (44%) yielded more grain yield than the normal check Based on the initial evaluations, the better performing “enhanced” lines were selected and intercrossed. Fifteen “enhanced B73 lines” were chosen for crossing with 18 “enhanced Mo17 lines” producing 93 hybrids that were evaluated in replicated trials. These “enhanced B73 X enhanced Mo17 hybrids” were evaluated at two planting densities and were compared with two checks (normal B73 x Mo17 and a high yielding Pioneer Commercial Hybrid 3165). Of the 93 hybrids, 6 hybrids exceeded the checks by 2 SD or more and two of the
six exceeded the checks by more than 15 per cent. These results demonstrated that marker assisted backcrossing can be successfully employed to manipulate complexly inherited traits such as grain yield.

**Introgression breeding**

Introgression breeding is used to transfer the favourable genes from the wild relatives to the cultivated species. The classical example of introgression breeding involving wide hybridization and mutation approach dates back to 1950s. Tanksley et al. (1981) used isozyme markers for selection backcross populations of tomato. However, the progress in this area is somewhat slow because of various problems associated with wide hybridization and the lack of an easy and reliable detection system to identify the introgressants. Cytological approaches were adapted to identify the lines having the introgressed segments, but the tediousness in the method made the process less amenable. Further, transferring specific genes from wild species to cultivated species is difficult, unless specific cytological markers are associated with those genes. With this approach one cannot aim to introgress favourable genes for agronomically important traits such as yield, nutritional quality, stress tolerance etc., since the traits are quantitatively inherited.

The ability to manipulate genes responsible for quantitative traits is a prerequisite for sustained improvement of crop plants. With the availability of recent molecular marker based QTL analysis, breeders can make their attempt to transfer the favourable QTLs from wild species to cultivated varieties. Tanksley and Nelson (1996) proposed an elegant method for the discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines and is termed as Advanced Back Cross-QTL analysis (AB-QTL analysis).

**AB-QTL analysis**

In attempting to transfer the favourable QTLs from unadapted germplasm to cultivated species, the regular method of QTL analysis using the balanced population cannot be used. The reasons are as follows. 1) influence of undesirable QTL alleles from the unadapted parent, 2) difficulty in detecting QTLs without epistatic interaction and 3) inability to detect subtle pleiotropic effects. A possible solution to the above problems would be to delay QTL analysis until an advanced generation (e.g. BC2 BC3 etc.). By this, 1) undesirable QTL alleles are eliminated, 2) QTLs with epistatic interactions are eliminated, 3) the deleterious effects due to linkage drags are less likely to be observed and 4) the process helps to construct QTL-NILS (genotypes similar to the recurrent parent except for specific QTLs. The steps involved in AB-QTL analysis for hybrid crops and various aspects on the use of AB-QTL analysis are described by Tanksley and Nelson (1996). Xiao et al (1996) recently used this approach to identify two genes from *Oryza rufipogon* which may have the potential to substantially increase the maximum yield potential of rice.

**Conclusion**

Many complex traits are under the control of multiple genetic factors, the effects of which are influenced by the environment. Because of these multiple influences, mapping the genetic basis (quantitative trait loci or QTL) for the variation in continuous traits is seldom easy. The advent of molecular marker technology and its potential use in QTL mapping made several aspects of crop breeding – marker aided selection for complex
traits, marker aided introgression of favourable QTLs from unadapted germplasm and tracing the orthology among crop groups – into reality. Thoday’s (1961) vision of making a saturated linkage map is made possible by the innovative molecular approach and in the same manner; one can expect to make the present day visions of crop breeding as future realities QTL, otherwise described as hypothetical genes based on statistical inferences, have very little biological meaning. To date the knowledge on QTL mapping is enormous. Issues related to genetic mapping of QTL are well reviewed by Liu (1998), Lynch and Walsh (1997) and Paterson (1998). Mauricio (2001) and Doerge (2002) critically reviewed on various approaches adopted for QTL mapping and analysis of QTL in experimental populations. Price (2006) elaborated the success stories of QTL mapping in his article “Believe it or not, QTLs are accurate”. However, the accrued knowledge does not have immediate solutions to the problems associated with QTL mapping. Detecting QTLs in biological systems has become a routine affair to make a way to understand the genetic architecture of complex traits. Still there is a kind of mistism in QTL mapping to exploit the potential in practical plant breeding (Fig 7).

Fig 7. Possible mistism around QTL mapping

Arunachalam and Chandrasekaran (1993) made a critical analysis on the historical papers on QTL mapping in tomato and reported that the experiments were conducted with small segregating populations, possibilities for biased association of markers with phenotypes, possibilities of biased threshold statistic to fix the QTLs resulting in false positives and non-detection of QTLs with minor effects. Flint and Mott (2001) listed the following pitfalls involved in QTL mapping.

1. There is no guarantee a QTL detected in one cross can be detected in another set of cross of the same species.
2. QTL detection experiments do not map a gene, but rather a genetic effect that might consist of many linked genes.
3. Most of the QTL detection experiments are not strong enough in elucidating interaction between loci.
4. Detection of large effect QTL does not harbour always a single large effect gene making the fine mapping of QTL difficult.

5. Fine mapping q QTL of moderate to large effect needs a very large population of recombinants. But large effect QTLs are proving to be rare.

The discussion on the possibilities of overcoming the pitfalls and deriving the biological meaning of QTLs is beyond the scope of this review. Though there are success stories of moving from QTL to gene and cloning the genes, the experimentation of QTL detection is associated with many shortcomings and many times these shortcomings are overlooked by the experimenters. Steinmetz et al. (2002) reported based on their QTL mapping work in yeast that the current techniques are likely to miss QTL if their effects are too small or if they are produced in part by unexpected combinations of alleles between stains. When this happens in simple organism like yeast, the plant scientists should move further with caution in exploring QTLs for complex traits in plants and plant geneticists should not assume QTL mapping is just a “cakewalk”.

References and suggested readings


