Electronic Journal of Biotechnology ISSN: 0717-3458
http://www.ejbiotechnology.info
DOI: 10.2225/vol13-issue5-fulltext-21

The genetic dissection of quantitative traits in crops

Kassa Semagn[1] ✉ · **Åsmund Bjørnstad**[2] · **Yunbi Xu**[3]

1 International Maize and Wheat Improvement Center, Nairobi, Kenya

2 Department of Plant and Environmental Sciences, Norwegian University of Life Sciences, Ås, Norway

3 International Maize and Wheat Improvement Center, Mexico D.F., Mexico

✉ Corresponding author: k.semagn@cgiar.org
Received June 2, 2009 / Accepted July 29, 2010
Published online: September 15, 2010

Abstract Most traits of interest in plant breeding show quantitative inheritance, which complicate the breeding process since phenotypic performances only partially reflects the genetic values of individuals. The genetic variation of a quantitative trait is assumed to be controlled by the collective effects of quantitative trait loci (QTLs), epistasis (interaction between QTLs), the environment, and interaction between QTL and environment. Exploiting molecular markers in breeding involve finding a subset of markers associated with one or more QTLs that regulate the expression of complex traits. Many QTL mapping studies conducted in the last two decades identified QTLs that generally explained a significant proportion of the phenotypic variance, and therefore, gave rise to an optimistic assessment of the prospects of markers assisted selection. Linkage analysis and association mapping are the two most commonly used methods for QTL mapping. This review provides an overview of the two QTL mapping methods, including mapping population type and size, phenotypic evaluation of the population, molecular profiling of either the entire or a subset of the population, marker-trait association analysis using different statistical methods and software as well as the future prospects of using markers in crop improvement.

Keywords: association mapping, linkage disequilibrium, markers assisted selection, molecular breeding, molecular markers, quantitative trait, QTL mapping, QTL analysis

INTRODUCTION

Plant breeding is a three step process, wherein populations or germplasm collections with useful genetic variation are created or assembled, individuals with superior phenotypes are identified, and improved cultivars are developed from selected individuals (Moose and Mumm, 2008). Figure 1 summarizes the different breeding methods that are commonly employed in crop improvement programs. Most of the traits of interest in plant breeding (*e.g.*, yield, height, drought resistance, disease resistance in many species, etc.) are quantitative, also called polygenic, continuous, multifactorial or complex traits. A quantitative trait is a measurable trait that depends on the cumulative action of many genes and their interaction with the environment

that can vary among individuals over a given range to produce a continuous distribution of phenotypes (Sham et al. 2002). Since the proposal of the multiple-factor hypothesis by Nillson-Ehle (1909) and East (1916), the genetic variation of a quantitative trait is assumed to be controlled by the collective effects of numerous genes, known as quantitative trait loci (QTLs) (Bulmer, 1985; Edwards et al. 1987; Falconer and Mackay, 1996; Xu, 1997; Lynch and Walsh, 1998; Xu, 2010). Consequently, several QTLs regulate the expression of a single phenotypic trait (in this paper, QTL refers to a single region of DNA associated with a particular trait while QTLs refers to the situation when two or more regions of DNA from the same or different chromosomes are associated with a particular trait).

Unlike monogenic traits, polygenic traits do not follow patterns of Mendelian inheritance (qualitative traits). Instead, their phenotypes typically vary along a continuous gradient depicted by a bell curve. Quantitative traits complicate the works of breeders because performance only partially reflects the genetic values of the individuals. If fruit size, for example, is controlled by a single gene with alleles “s” for small and “S” for large, then the progeny of crosses between the two parents would segregate in to 3:1 ratios of large- to small- fruited plants. For such discrete traits, one can infer the "genotype" (SS or Ss versus ss) by observing the "phenotype" (large or small). For quantitative traits, the situation is more complex: (i) quantitative traits are controlled by multiple genes or QTLs, and plants with the same phenotype can carry different alleles at each of many genes or QTLs; (ii) plants with identical QTL genotypes can show different phenotypes when raised under different environments; and (iii) the effect of one QTL can depend on the allelic constitution of the plant at other QTL. For these reasons, one cannot infer the genotype from the phenotype, and one must construct specialized genetic stocks and grow them in precisely controlled environments.

QTLs have been identified for quantitative traits as reported in the literature. The number of QTLs detected in a given study depends on different factors, including type and size of mapping population used, trait investigated, the number of environments used for phenotyping, and genome coverage. The QTLs reported in the literature include two groups of genes. The first group constitutes a small proportion of the published literature and includes major genes of very large effects on highly heritable traits, with each explaining a large portion of the total trait variation in a mapping population. Most QTLs reported in the literature fall in another group that are regulated by many genes, each explaining small portion of the total trait variation. For example, Laurie et al. (2004) reported about 50 QTLs that explained approximately 50% of the genetic variance for oil concentration in the maize kernel. Buckler et al. (2009) evaluated nearly a million maize plants in eight environments and found no evidence for any single large effect QTL for flowering time. The authors identified numerous QTLs of small additive effects that are shared among families. However, the genetic variation of most quantitative traits likely involves a small number of major genes or QTLs, a larger number of loci with moderate effects, and a very large number of loci with minor effects (Robertson, 1967; Kearsey and Farquhar, 1998). The effects of the major genes can be studied via segregation analysis as well as evolutionary and selection history. The numerous genes with small effects, however, cannot be investigated individually.

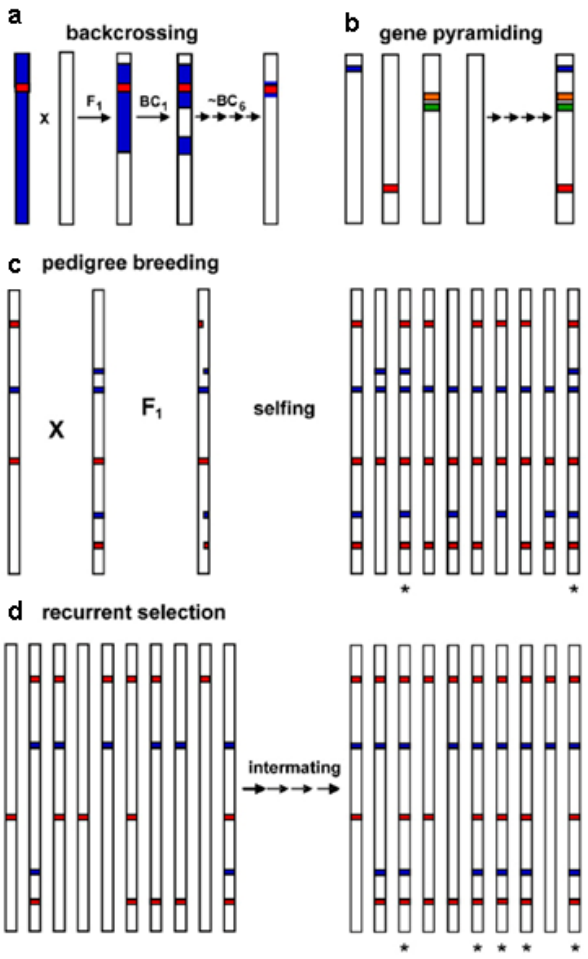

Fig. 1 Common plant breeding and selection schemes. Each vertical bar is a graphical representation of a chromosome of an individual within a breeding population, with colored segments indicating genes and/or QTLs that influence traits under selection. Genes associated with different traits are shown in different colors (*e.g.* red, blue). ''X'' indicates a cross between parents, and arrows depict successive crosses of the same type. Asterisk below an individual signifies a desirable genotype. (a) In backcrossing scheme, a donor line (blue bar) is crossed to an elite line for transferring a specific gene of interest (red). Selected progenies were repeatedly backcrossed to the elite parent with each backcross cycle involving selection of individuals with the gene of interest and of the highest proportion of elite parent outside the target genome. (b) In gene pyramiding, genes or QTLs associated with different beneficial traits (blue, red, orange, green) are combined into the same genotype via crossing and selection. (c) In pedigree breeding, two individuals with desirable and complementary phenotypes are crossed; F_1 progeny are self-pollinated to fix new, improved genotype combinations. (d) In recurrent selection, a population of individuals (10 in this example) segregate for two traits (red, blue), each of which is influenced by two major favorable QTLs. Intermating among individuals and selection for desirable phenotypes/genotypes increases the frequencies of favorable alleles at each locus. For this example, no individual in the initial population had all of the favorable alleles, but after recurrent selection half of the population possesses the desired genotype (Moose and Mumm, 2008).

The theory of QTL mapping was first described by Sax (1923), where he noted that seed size in bean (a complex trait) was associated with seed coat color (a simple, monogenic trait). This concept was further elaborated by Thoday (1961), who suggested that if the segregation of simply inherited monogenes could be used to detect linked QTLs, then it should eventually be possible to map and characterize all QTLs involved in complex traits. Before the advent of modern QTL mapping, traits showing quantitative variation were studied by statistical analysis of appropriate experimental populations based on the means, variances and covariances of relatives, with no actual knowledge of the number and location of the genes that underlie them (Kearsey and Farquhar, 1998). These studies focused on phenotypic distributions of populations and correlations in phenotypes among related individuals or lines. New interest in QTL mapping in crops was generated when studies on fruit traits of tomato (Paterson et al. 1988) and the morphological and agronomic characters of maize (Stuber et al. 1992) successfully demonstrated that some molecular markers explained a substantial proportion of the phenotypic variance of quantitative traits.

The two general goals of QTL mapping in plants are to (a) increase our biological knowledge of the inheritance and genetic architecture of quantitative traits, both within a species and across related species, and (b) identify markers that can be used as indirect selection tools in breeding (Bernardo, 2008). During the past two decades, the ability to transfer target genomic regions using molecular markers resulted in extensive QTL mapping experiments in most economically important crops, aiming at the development of molecular markers for marker assisted selection (Xu, 1998; Collard et al. 2005; Semagn et al. 2006a; Xu, 2010) and QTL cloning (Salvi and Tuberosa, 2005). Results from such studies provide information on (a) the number and chromosomal location of QTLs affecting a trait; (b) the magnitude and direction of effect of each QTL (*i.e*., whether a phenotypic trait is controlled by many genes or many independent loci of small effect or by a few genes of large effect); (c) the mode of gene action at each QTL (dominant or additive); (d) the parental sources of beneficial QTL alleles, and (e) whether there is interaction between different QTLs (epistasis, *i.e.,* interactions between two QTLs that result in an effect on the trait that would not be predicted from the sum of the individual QTL effects) or between genotypes and environment (Bradshaw, 1996). Figure 2 and Table 1 summarizes results of QTL mapping study in a double haploid hexaploid wheat population for Fusarium head blight resistance, deoxynivalenol content and anther extrusion. Most studies identified QTLs that generally explained a significant proportion of the phenotypic variance of the respective trait, and therefore, gave rise to an optimistic assessment of the prospects of markers assisted selection. However, several studies reported many QTLs for a given trait with multiple QTLs in every chromosome. Few examples include yield in maize (Tuberosa et al. 2002), nematode resistance in soybean (Concibido et al. 2004) and Fusarium head blight resistance in hexaploid wheat (Kolb et al. 2001).

QTL mapping requires that the researcher (1) select and/or develop appropriate mapping population (experimental populations for linkage-based mapping or natural/breeding populations for association mapping); (2) phenotype the population for the trait(s) of interest (morphological characters, agronomic traits, disease and pest scores, drought resistance, etc.) under greenhouse, screen-house and/or field conditions; (3) decide the type of molecular marker(s), the genotyping approach (entire population, selective genotyping or bulk segregant analysis) and generate the molecular data for adequate number of uniformly-spaced polymorphic markers; (4) identify molecular markers linked to the trait(s) of interest using statistical programs

(linkage-based QTL mapping methods requires construction of genetic linkage map); and (5) test the applicability and reliability of the markers associated with major QTLs in predicting the trait(s) in related families (marker validation or verification) for QTLs of medium to large effect. Details on molecular markers, genetic linkage mapping and marker assisted selection have been previously reviewed by Semagn et al. (2006a), Semagn et al. (2006b) and Semagn et al. (2006c). The availability of a wide range of molecular markers and powerful statistical methods has significantly facilitated QTL mapping (Figure 3, Figure 4). Linkage analysis and association mapping are the two most commonly used tools for dissecting complex traits. Both QTL mapping methods begins with the collection of genotypic and phenotypic data from either segregating or natural population, followed by statistical analyses to reveal all possible marker loci where allelic variation correlates with the phenotype. This article provides an overview of the various issues related to the two QTL mapping methods and their future prospects in crop improvement programs. Jansen and Nap (2001) introduced the concept of genetical genomics, in which genetics and gene expression approaches have been joined in detecting expression quantitative trait loci (eQTL) that control the observed variation in gene expression. eQTL mapping is different from QTL mapping, since researchers are mainly interested in major eQTL in cis (within the gene) and major regulatory eQTL in trans. However, eQTL mapping is not part of this review paper.

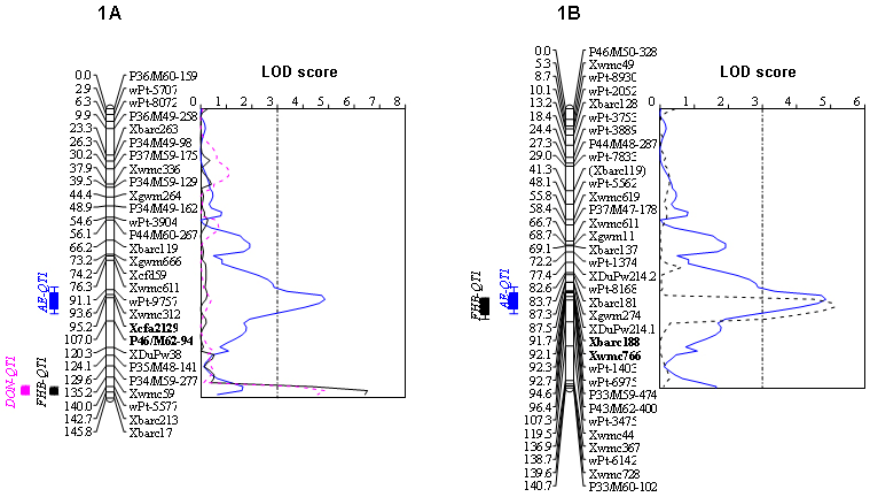

Fig. 2 A likelihood-ratio test statistic (LOD score) plot of hexaploid wheat chromosomes 1A and 1B showing QTLs for Fusarium head blight (FHB) resistance (black dotted line), anther extrusion (AE, solid blue line) and deoxynivalenol (DON) accumulation (pink dashed line). LOD scores were obtained for mean phenotypic dataset using composite interval mapping. QTL graphs and a threshold of LOD 3.0 are shown on the right of each chromosome. Genetic distances are shown in centimorgans (cM) on the left of each chromosome. The flanking markers for anther extrusion as an example are shown in bold face (Semagn et al. 2007; Skinnes et al. 2010).

TYPES OF MAPPING POPULATIONS

Choice of appropriate mapping population is very critical for the success of any QTL mapping project. Populations for QTL mapping can be broadly classified into two: experimental populations for linkage-based QTL mapping (*e.g.,* inbred lines for autogamous or self pollinating species; half- or full- sib families for outcrossing or cross pollinating species) and natural or breeding populations for linkage disequilibrium-based association mapping. For association mapping, the populations can be classified into one of the following five groups (Yu and Buckler, 2006; Yu et al. 2006): (i) ideal sample with subtle population structure and familial relatedness, (ii) multi-family sample, (iii) sample with population structure, (iv) sample with both population structure and familial relationships, and (v) sample with severe population structure and familial relationships. Due to local adaptation, selection, and breeding history in many plant species, many populations for association mapping would fall into category four (Zhu et al. 2008). Alternatively, populations for association mapping can be classified according to the source of materials as germplasm bank collections, synthetic populations, and elite germplasm (Breseghello and Sorrells, 2006).

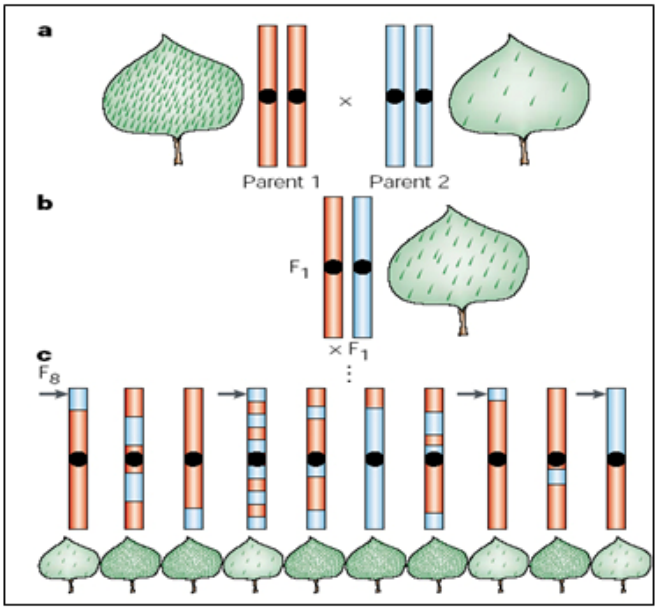

Fig. 3 Principles of mapping quantitative trait loci (QTL). (a) Inbred parents that differ in the density of trichomes (parent 1: high trichome density; Parent 2: low trichome density) are crossed to form an F_1 population with intermediate trichome density. (b) An F_1 individual is selfed to form a population of F_2 individuals. (c) Each F_2 is selfed for six additional generations, ultimately forming a set of recombinant inbred lines (RILs). Each RIL is homozygous for a section of a parental chromosome. The RILs are scored for genetic markers, as well as for the trichome density phenotype. In (**c**), the arrow marks a section of chromosome that derives from Parent 2 (the parent with low trichome density). The leaves of all individuals that have inherited that section of chromosome from the parent with low trichome density also have low trichome density, indicating that this chromosomal region probably contains a QTL for this trait (Mauricio, 2001).

Linkage-based QTL mapping depends on well defined populations developed by crossing two parents. In autogamous species, QTL mapping studies make use of F_2 or F_x derived families, backcross (BC), recombinant inbred lines (RILs), near isogenic lines (NILs), and double haploids (DH). These populations are developed by crossing two inbred parents with clear contrasting difference in phenotypic trait(s) of interest. Each mapping population developed from inbred parents has its own advantages and disadvantages and the researchers need to decide the appropriate population depending on project objective, trait complexity, available time, and whether the molecular markers to be used for genotyping are dominant or codominant. Both F_2 and BC populations are the simplest types of mapping populations because they are easy to construct and require only a short time to produce. F_2 is more powerful for detecting QTLs with additive effects, and can also be used to estimate the degree of dominance for detected QTLs. When dominance is present, backcrosses give biased estimates of the effects because additive and dominant effects are completely confounded in this design (Carbonell et al. 1993). However, both F_2 and BC populations have three limitations. First, development of these populations require relatively few meioses such that even markers that are far from the QTLs remain strongly associated with it. Such long-distance associations hamper precise localization of the QTLs. Second, F_2 and backcross populations are temporary populations as they are highly heterozygous and cannot be propagated indefinitely through seeds (*i.e*., these populations can't be evaluated several times in different environmental conditions, years, locations, etc.). Finally, epistatic interactions could hardly be studied in both F_2 and backcross populations.

In classical quantitative genetics, if a trait has a low heritability, one can take the family mean as the unit of measurement and select the parents with high average performance on the basis of the family mean (Mather and Jinks, 1982) because family-mean-based heritability can be significantly increased by increasing the number of progenies. This idea has been first applied to genetic mapping for low heritability traits in animals by using the daughter or granddaughter designs, where the phenotypic value of the sire has been replaced by the mean phenotypic value of the daughters (Weller et al. 1990; Ron et al. 2001). The same idea was then applied to plants by replacing the phenotypic value of an F_2 plant by the mean of F_3 progeny, called the $F_{2:3}$ design (Austin and Lee, 1996; Fisch et al. 1996). All F_3 progeny derived from the same F_2 plant belong to the same $F_{2:3}$ family, denoted by $F_{2:3}$. If the size of each $F_{2:3}$ family (the number of F_3 progeny) is sufficiently large, the average value of the family will represent the genotypic value of the F_2 plant, and thus the power of QTL mapping may significantly increase. One can increase the number of generations from 3 to y leading to an $F_{x:y}$ design. In such cases, genotyping will be done on individuals plants in generation x and phenotyping in generation y with $y > x$ (Fisch et al. 1996; Jiang and Zeng, 1997; Chapman et al. 2003). Alternatively, genotyping can also be done by bulking DNA or leaf tissue of at least 15 individuals from the same family at generation y. As y increases at least to 6 generations, the design becomes the RILs design. RILs are derived from an F_2 population by generations of full-sib mating (mating between offspring's from the same parents for outcrossing species) or selfing (bulk or single seed descent) (Soller and Beckman, 1990; Xu and Crouch, 2008). RILs are advanced homozygous lines that have undergone several rounds of inbreeding (Darvasi and Soller, 1995). Such multiple generations of mating increases the potential number of recombination events and improves map resolution (*i.e*., sufficient meioses have occurred to reduce disequilibrium between moderately linked markers).

If backcross selection is repeated at least for six generations, more than 99% of the genome of randomly selected individuals at BC_6 and above will be derived from the recurrent parent. Selfing of selected individuals from BC_7F_1 will produce two types of BC_7F_2 lines that are homozygous for the two alleles at the target gene locus, which are said to be nearly isogenic with each other and with the recipient parent (NILs). Heterogeneous inbred family analysis was also proposed as a method to quickly develop NILs for an identified QTL in inbred lines (Harris et al. 2007; Pumphrey et al. 2007; Xu and Crouch, 2008). Selection for the target trait is required for the generation of NILs. By essentially fixing the genetic background, NILs are ideal for construction of high-resolution mapping, gene expression profiling, and more direct hypothesis-driven biological experimentation. NILs are particularly effective genetic stocks for studying phenotypic effects attributable to a QTL since the genetic background, including morphological and phenological characters that commonly influence phenotypic assessments of quantitative traits, is uniform. Double haploid (DH) populations have also been used for QTL mapping in several species (*e.g.*, Bao et al. 2002; Mahmood et al. 2003; Behn et al. 2005; Semagn et al. 2006d; Semagn et al. 2007; Xu and Crouch, 2008). The DH production methodology improves breeding efficiency by generating inbred lines with 100% purity and genetic uniformity in just two generations. DH lines make it easy to carry genetic studies and shorten the breeding time significantly.

Table 1. Summary of the composite interval mapping (CIM) analysis of quantitative trait loci (QTL) for mean Fusarium head blight severity, deoxynivalenol (DON) content and anther extrusion in a double haploid population derived from the cross between Arina and NK93604. For each QTL, chromosomal location, marker interval, LOD, percent of explained phenotypic variance (R^2), and the parent contributing the favored allele are listed. A QTL was declared significant at LOD ≥ 3.0 (modified from Semagn et al. 2007 and Skinnes et al. 2010).

Chromosome	Marker interval		Map position (cM)	Confidence interval (cM)	Fusarium head blight		Deoxynivalenol content		Anther extrusion		Parental source for beneficiary allele
	Left marker	Right marker			*LOD*	*R^2*	*LOD*	*R^2*	*LOD*	*R^2*	
1AL	Xcfa2129	P46/M62-94	96	92-102					4.9	18.3	NK93604
1AL	wPt-5577	Xbarc213	142	140-144	6.5	27.9	5.9	27.9			NK93604
1BL	Xbarc188	wmc766	92	86-102					3.2	7.4	Arina
1BL	P43/M62-400	wPt-3475	100	94-104	5.1	19.6	-	-			Arina
2AS	wPt-6148	Xbarc124.1	16	14-18	-	-	5.8	26.7			NK93604
4DL	XDupW278	Xgwm624	54	46-62					4.6	13.3	Arina
6AS	P33/M50-257	Xbarc3	6	4-10					6.0	15.6	Arina
6BS	P46/M62-107	P45/M60-265	68	64-72	3.0	7.8	-	-			Arina
7AL	Xgwm276	XDuPw226	100	94-106	4.9	14.8	-	-			NK93604
	Simultaneous fit (adjusted R^2)				13.9	49.1	8.4	34.0	15.7	53.6	

RILs, NILs and DHs are permanent populations because they are homozygous or 'true-breeding' lines that can be multiplied and reproduced without genetic change occurring. Seeds from RILs, NILs and DHs can be transferred between different laboratories for mapping to ensure that all collaborators examine identical material (Young, 1994; Paterson, 1996; He et al. 2001) so that genetic results from phenotyping, genotyping and QTL mapping can be accumulated across laboratories. The main limitations of NIL and RIL include (i) the long time and/or high cost required to develop these populations, and (ii) these populations only detect the additive

component but provide no information on dominance relationships for any QTL (Haley and Andersson, 1997). DH populations are quicker to generate than RILs and NILs but the production of DHs is only possible for species with a well established protocol for haploid production. The limitations common to all mapping populations developed from inbred lines include (a) the confidence interval for many QTLs mapped using the most commonly used population size (100-200 samples) is several centimorgans (abbreviated as cM), which could correspond to hundreds of genes (Kroymann and Mitchell-Olds, 2005); (b) the low number of alleles sampled per locus in each population (Figure 5) makes it difficult to examine the full range of genetic diversity available for many plant species; and (c) for some species such as outcrossing, it is often impossible due to inbreeding depression or self incompatibility or very impractical, time consuming and/or expensive to produce inbred lines.

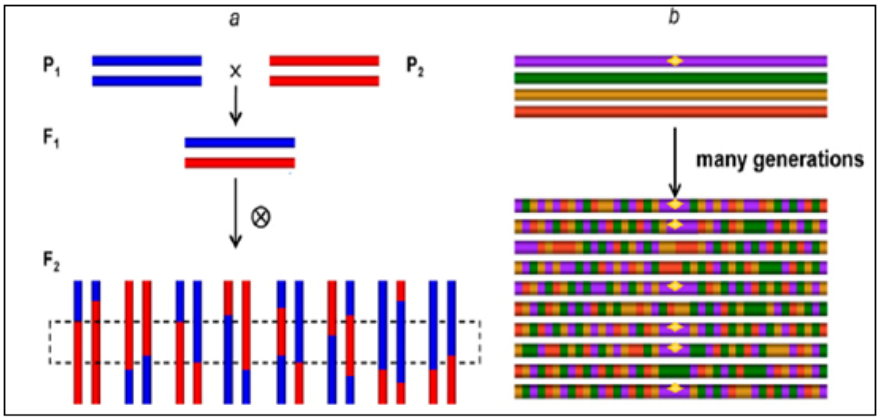

Fig. 4 Schematic comparison of linkage analysis with designed mapping populations and association mapping with diverse collections. In linkage analysis (panel a, using F_2 design as an example), there are only few opportunities for recombination to occur within families and pedigrees of known ancestry, resulting in relatively low mapping resolution). In association mapping (panel b haplotype) historical recombination and natural genetic diversity were exploited for high resolution mapping. Linkage disequilibrium between a functional locus (yellow diamond for mutated allele) and molecular markers is low except for those within very short distance (Zhu et al. 2008).

Genetic analyses in outcrossing species are far more complicated than species that can be selfed to produce inbred lines. Some of the difficulties arise when heterozygous and heterogeneous parents are crossed to develop a mapping population. First, the number of marker alleles and the segregation pattern of marker genotypes may vary from locus to locus in outcrossing species, whereas an inbred line-initiated segregating population always has two alleles and an expected segregation ratio across different markers. Second, complications arise if parents have alleles in common at the QTL or marker loci, or if the parents share QTL alleles in different linkage phases with the marker loci (Jansen et al. 1998; Lynch and Walsh, 1998). Third, linkage phases among different markers are not known a priori for outbred parents and, therefore, an algorithm should be used to characterize a most likely linkage phase for linkage analysis (Lu et al. 2004). To overcome these problems, other strategies based on two-way pseudo-testcross, half-sib and full-sib families derived from controlled crosses have been proposed for outcrossing species

(Knott and Haley, 1992; Mackinnon and Weller, 1995; Hoeschele et al. 1997; Uimari and Hoeschele, 1997; Liu and Dekkers, 1998; Xu, 1998; Sillanpää and Arjas, 1999). Grattapaglia and Sederoff (1994) proposed a two-way pseudo-testcross mapping strategy in which one parent is heterozygous whereas the other is null for all markers. Using this strategy, two parent-specific linkage maps will be constructed. The limitation of the pseudo-testcross strategy is that it can only make use of a portion of molecular markers. Several other authors (*e.g*., Ritter and Salamini, 1996; Maliepaard et al. 1998; Wu et al*.* 2002) proposed various approaches for determining the linkage and parental linkage phases for any type of molecular markers. Ma et al*.* (2004) devised a general model for estimating the probability of parental linkage phases, which allows for a simultaneous estimation of the linkage. Wu et al*.* (2002) and Lu et al. (2004) constructed a unifying likelihood analysis to simultaneously estimate linkage, linkage phases and gene order for a group of markers that display all possible segregation patterns in a full-sib family derived from two outbred parents.

Population size and environment

There is usually a high cost associated with genotyping (generation of molecular marker data) and phenotyping (field, greenhouse or screen house evaluation for the phenotypic trait) of large population size, particularly for traits requiring extensive field trials or complex analysis. Consequently, the size of the mapping population and the number of replications and sites (environments) for phenotyping is often limited. Thus, most published experiments with replicated trials have used between 100 and 200 progenies (*e.g*., Lynch and Walsh, 1998; Somers et al. 2003). Overall, the QTL mapping literature has shown that if a breeder can develop a mapping population of 100-150 progenies derived from an F_2 or backcross population between two inbreds, obtain reasonably good phenotypic data for the traits of interest, and genotype the population with markers spaced about 10 to 15 cM apart, then an analysis of the phenotypic and marker data with an appropriate statistical method will almost always lead to the identification of at least a few markers associated with each trait of interest (Bernardo, 2008). However, small population size often resulted in the detection of few QTLs with large phenotypic effects (Beavis, 1998; Melchinger et al*.* 1998; Utz et al*.* 2000; Schon et al. 2004). Nonetheless, it does not necessarily indicate that QTL position will be inaccurate although this may be the case.

Melchinger et al. (1998) evaluated the power of QTL detection of different traits in maize by comparing results from QTL mapping in two independent samples of different size from the same population (344 F_2 population in experiment-1 and 107 F_2 population in experiment-2). The total number of QTLs detected for all traits in experiment-1 was almost triple to that of the numbers detected in experiment-2. Only about half of the putative QTLs detected in experiment-2 were in common with QTLs identified in experiment-1. In addition, the magnitude of QTL effects can also be biased by small sample size. In a study on QTL experiments in maize, for example, Beavis (1998) identified one or a few QTLs of large effect along with several QTLs of small effect. The fewer the progeny, the higher were the estimated effects of the largest QTLs identified. Similarly, Melchinger et al. (2004) partitioned their entire dataset for maize testcross progenies (N = 976 genotypes and E = 16 environments) into smaller datasets (N = 488, 244, 122 and E = 16, 4, 2) and clearly demonstrated highly inflated QTL effects for the smaller samples. Furthermore, the QTLs of large phenotypic effect can also be an artifact of the strong directional selection often used to create the phenotypically divergent parental lines that are used for mapping (Lande, 1983).

One of the first published reports of QTL mapping in crops that utilized molecular markers involved fruit size, pH, and soluble solids in tomato (Paterson et al. 1988). A total of 237 backcross progenies from a cross between cultivated tomato (*Lycopersicon esculentum)* and a wild relative (*L. chmielewskii*) were analyzed in one location and genotyped with 70 restriction fragment length polymorphism (RFLP) markers. That study detected 6 QTLs for fruit size, 5 QTLs for pH, and 4 QTLs for soluble solids. Subsequently, an F_2 population of 350 individuals derived from a cross between *L. esculentum* x *L. cheesmanii*, along with corresponding F_2-derived F_3 families was analyzed in three locations (environments). A total of 29 QTLs for fruit size, pH, and soluble solids were detected. However, only 4 QTLs were consistently detected in all three environments with 10 QTLs in two environments and 15 QTLs in only one environment (Paterson et al. 1991). QTLs of both major (as high as 40% of total variation) and minor (as little as 4%) effects were found for all traits. Altogether, the identified QTLs for fruit size accounted for 76% of total variation in the trait, 44% of total variation in soluble solids, and 34% of total variation in fruit pH. The remainder of the variation was presumably a result of (a) environment, (b) measurement error, (c) additional QTLs with effects too small to be detected with confidence in such population size, (d) interactions between QTLs, which were too small to detect, and (e) genotype-by-environment (GxE) interactions (Young, 1996).

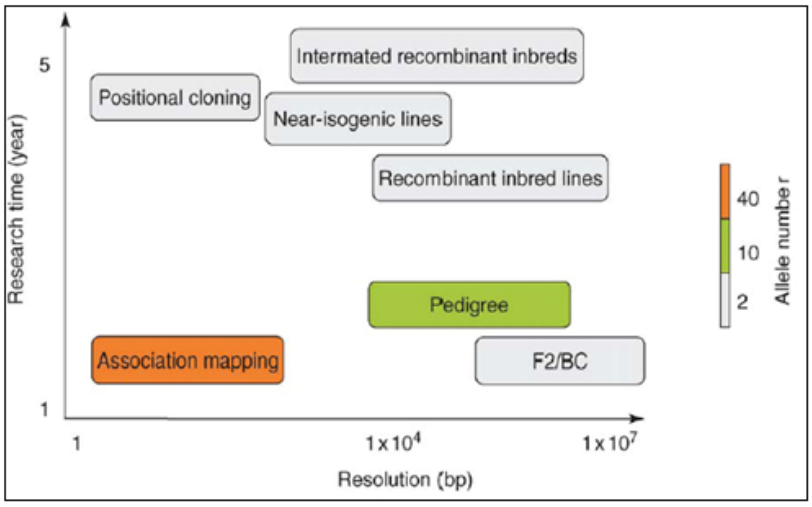

Fig. 5 Schematic comparison of various methods for identifying marker-trait association in terms of resolution, research time and allele number. F2/BC refers to F_2 and backcross populations (Yu and Buckler, 2006).

Schon et al. (2004) used a dataset composed of 976 F_5 maize testcross progenies evaluated in 19 environments and cross-validation to assess the effect of sample size (N), number of test environments (E), and significance threshold on the number of detected QTL, the proportion of the genotypic variance explained by them, and the corresponding bias of estimates for grain yield, grain moisture, and plant height. The number of detected QTLs and the proportion of genotypic variance explained by QTLs generally increased more with increasing N than with increasing E. The average bias

of QTL estimates and its range were reduced by increasing N and E. A substantial bias was found for estimates of the proportion of genotypic variance explained by the detected QTLs even with N = 976, irrespective of the trait, the heritability, and the significance threshold. This confirm results from the study by Beavis (1998), who pointed out that the bias of QTL estimates could not be ignored even for N > 500. As pointed out by different investigators (*e.g.*, Knapp et al. 1990; Moreau et al. 1998), it is therefore advisable to increase population size rather than the number of test environments or replications for most traits unless plot heritabilities are very low and/or the expenditures for molecular analyses of additional genotypes are much higher than those for additional testing of phenotypes. The comparison of subpopulations with the same plot capacities for phenotypic evaluation revealed that increasing the number of progenies generally increased the power of QTL detection and the proportion of the genotypic variance explained by QTL and reduced the bias more efficiently than did increasing the number of test environments. Although decision regarding population size and number of phenotyping environments depend on several factors (*e.g.*, capacity and resource availability, population type, trait heritability, marker type for genotyping, ease in phenotyping, etc.), we recommend at least 184 progenies and 3 phenotyping locations (environments). This number is recommended based on our experience and enables to organize the entire mapping population in two 96-wells plates (each consisting of 92 progenies, 2 parents, F_1 and a negative control) either for in-house genotyping or outsourcing (genotyping by service providers).

Phenotyping

The basic phenotypic data required for QTL mapping are the estimates of phenotypic performance of individuals across environments. The accuracy and precision of phenotyping determines how realistic the QTL mapping results are. The power of QTL detection, defined as the probability of detecting a QTL at a given level of statistical significance (Manly and Olson, 1999), depends upon the number of progeny in the population (sample size), heritability of the trait, genetic dissimilarity among progenies, the effect of the QTLs, and the environment used for phenotypic evaluation. Due to the availability of high-throughput and low cost molecular tools, genotyping no longer limit the sample size in mapping studies but the cost and logistics of phenotyping impose limits on sample size. This is especially true of phenotypes involving complex traits (Jin et al. 2004). The level of heritability of a trait depends in part on whether the phenotyping is repeatable across different seasons, locations and environments. Increased precision of phenotyping increases heritability which, in turn, increases the statistical power of QTL detection.

An appropriate phenotyping protocol should consist of (a) a representative sample of environments and their optimal location; (b) number of replications per individual in each environment; (c) experimental design to effectively account for extraneous variation in experimental field; (d) appropriate statistical methods for efficient analysis of data; and (e) consideration of QTL x environmental interaction. Replication and randomization of individuals and local control of errors, when properly used, have three benefits: (a) they allow separation of signal (the true differences in phenotypic performance among individuals) from noise; (b) they maximize the signal-to-noise ratio; and (c) they deliver a valid and unbiased estimation of level of noise/uncertainty in results. Replication simply indicates the number of plots assigned to an individual. It is necessary to obtain an internal estimate of experimental error variance and to permit separation of the genotype-environment interaction error variance.

Randomization provides statistical validity to results and protection from bias. Local control of error can be achieved by proper blocking of plots in a manner that maximizes inter-block and minimizes intra-block variation. Orientation of the blocks, as far as possible, should be perpendicular to the expected gradient in the experimental field, glasshouse bench, etc. However, there is always some variation left uncontrolled within blocks. Cross-population and environment comparison of phenotyping is needed in order to determine how the marker-trait association identified under one environment can be used for selection under another (Xu and Crouch, 2008).

Genotyping

Genotype (molecular markers) data can be generated in either of the following three ways: (1) by genotyping an entire mapping population; (2) by genotyping part of the population that exhibit extreme phenotypes for the target trait, known as selective genotyping (Lander and Botstein 1989; Darvasi, 1997; Vision et al. 2000; Micic et al. 2005; Xu and Crouch, 2008); or (3) by genotyping bulks of selected individuals, known as bulk segregant analysis (Giovanoni et al. 1991; Michelmore et al. 1991; Perez-Enciso, 1998; Breen et al. 1999; Fu, 2003). The usual QTL mapping method requires genotyping an entire mapping population with markers distributed across the whole genome. Such approach is more reliable but extensive, time consuming and expensive (Xu and Crouch, 2008). The second approach is selective genotyping (Figure 6) which involves genotyping of selected individuals that are chosen on the basis of the individuals' phenotypes (generally those with extremely high and/or low phenotypic values). Selective genotyping reduces the number of individuals that needs to be genotyped to detect QTLs by using only individuals at one or two extreme tails of the phenotypic distribution for the quantitative trait of interest (Lebowitz et al. 1987; Foolad et al. 1997; Prasad et al. 1999; Roy et al. 1999; Foolad et al. 2001; Ayoub and Mather, 2002; Zhang et al. 2003; Xu and Crouch, 2008). Selective genotyping is useful in situations in which full-population genotyping is too costly or not feasible, or where the objective is to rapidly screen large numbers of potential donors for useful alleles with large effects. Unidirectional selective genotyping (genotyping one side of the tail) is of particular interest for application within breeding programs, because it has the potential to permit QTL detection using superior progeny that have been retained under selection in breeding programs (Navabi et al. 2009). This allows larger numbers of potential donors to be screened for useful alleles with effects across different backgrounds.

There is no clear consensus regarding the number of individuals that need to be sampled from each tail. In a population of 436 recombinant inbred rice lines segregating for a large-effect QTLs affecting grain yield under drought stress, Navabi et al. (2009) reliably detected the QTLs by genotyping as few as 20 selected lines (4.5%). According to Ayoub and Mather (2002), genotyping of only 10% of the entire population was sufficient to detect all major QTLs. Darvasi and Soller (1992) showed that genotyping individuals only from the upper and lower 25% tails of the phenotypic distribution was nearly as efficient in detecting QTLs as genotyping the entire population. Gallais et al. (2007) suggested genotyping of about 30% for each tail. As the population size increases, the proportion of individuals required from each tail will decrease such that at a certain point an absolute number of plants from each tail will become the critical issue (Sun et al. 2010). However, selective genotyping has not been widely adopted, possibly due to distorted segregation in the production of linkage maps (Martinez, 1996), the biased estimates of the effects of linked QTLs (Lin

and Ritland, 1996), and the constraint of being able to study only a single trait at a time. Selective genotyping reduces the size of a mapping population that will, in general, decrease the power of QTL detection (Charcosset and Gallais, 1996), increase the QTL confidence interval, and increasing the probability of detecting false positive QTLs.

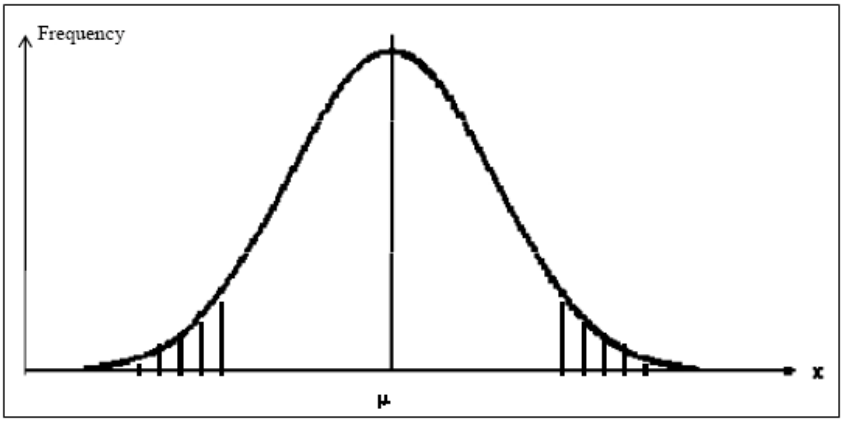

Figure 6. **Principle of selective genotyping and bulk segregant analysis.** The method of selective genotyping implies the selection of individuals with extremely high and low phenotypic values (within the shaded area) from the continuous distribution of a quantitative trait of interest. The selected individuals are genotyped and association is tested. Bulking segregant analysis advances selective genotyping approach one step further by bulking selected individuals at each of the two extreme trails, with each tail represented only by one bulk sample.

The third approach is the bulking strategy (Figure 6) that advances the selective genotyping approach one step further by using either plant bulking (bulking equal weight of leaf from each sample prior to DNA extraction) or DNA pooling (bulking DNA after extraction and normalization to the same concentration) from the selected individuals at each of the two extreme phenotypes. BSA measures the variation present in pools of segregants that have been sorted according to phenotype and uses the correlation between these measurements and the pool phenotype to assign a likely map location (Brauer et al. 2006). BSA has been successfully used in mapping single major genes (Barua et al. 1993; Villar et al. 1996) and two to three major QTLs (Quarrie et al. 1999; Shen et al. 2003) with a considerable research-saving compared to the previous two genotyping approaches. Theoretical analysis of BSA for experiments involving backcross, F_2 and half-sib designs shows that the power of selective DNA pooling for detecting genes with large effect can be the same as that obtained by individual selective genotyping. However, BSA is generally not regarded as a useful approach for either detection of QTLs which may be conditioned by several genes with small effect, or when the QTL is loosely linked to the marker. This is because the two bulks are frequently contaminated with alternative alleles if mischaracterization exists or recombination occurs (Darvasi and Soller, 1994; Wang and Paterson, 1994). As reviewed by Xu and Crouch (2008), the reliability of BSA for QTL mapping can be affected by (i) insufficient marker density; (ii) small population sizes, resulting in phenotypic differences between pools that are sufficient only to identify large-effect genes or QTLs; (iii) inaccurate estimate of allele frequencies

within pools; and (iv) high level of false positives. Sun et al. (2010) indicated that these problems can be solved by increasing population and tailed sizes and marker density.

QTL MAPPING

Having generated and entered both the phenotype and genotype data, researchers are often eager to test the two hypotheses in QTL analysis or QTL mapping: (i) null hypothesis (H_0): no QTL is present or a QTL is present but it is not linked to the marker(s) and (ii) alternative hypothesis (H_A): a QTL is present and it is linked to the marker(s). Various statistical methods exist for testing the two hypotheses (*e.g.*, Manly and Olson, 1999; Broman, 2001; Mauricio, 2001), which can be grouped into three based on the type of population(s) for mapping: (a) those methods that require the development of appropriate mapping population(s) using designed crosses (analysis of variance, simple interval mapping, composite interval mapping, multiple interval mapping); (b) those methods that use natural or breeding populations (*e.g.*, linkage disequilibrium-based mapping) and (iii) those methods that use either appropriate mapping populations or natural or breeding populations (*e.g.*, principal component analysis-based mapping and partial least square regression). The statistical methods for QTL mapping can also be grouped into two based on their requirements for genetic maps: (a) those methods that don't require prior genetic linkage map construction (analysis of variance, linkage disequilibrium-based mapping, principal component analysis-based mapping, partial least squares regression,) or (b) those that require availability of genetic map for the population (simple interval mapping, composite interval mapping, multiple interval mapping). For the latter, researchers need to conduct linkage analyses on the genotypic data and construct a genetic linkage map (Semagn et al. 2006d; Semagn et al. 2006a) for the population prior to QTL analysis. The statistical methods can also be grouped into two based on the distribution of phenotypic traits: (a) parametric methods (those that assume normal distribution) or require mathematical transformation of the phenotypic data into approximate normal distribution or (b) non-parametric (distribution free) methods. For a comprehensive coverage about the statistical methods for QTL detection, see Xu (2010). In this section, only basic statistical methods that have been used in QTL mapping will be described without providing details on statistical issues.

Linkage analysis-based QTL mapping

Statistical methods. Analysis of variance (ANOVA) is the simplest method for QTL mapping (Soller et al. 1976). Broman (2001) reviewed details of the methodology of QTL mapping using ANOVA. Once genotypic (molecular markers) and phenotypic (*e.g.*, disease scores, morphological characters, and agronomic traits) data are available for the population in question, ANOVA tests the statistical association of molecular markers to the phenotypic traits of interest. At each typed molecular marker, one splits the progenies into two groups, according to their genotypes at the marker, and compares the phenotype distributions of the two groups. The marker locus being tested on a given analysis is called the target locus. The test may include additional marker loci, called background markers, that have been shown to be associated with the trait and therefore lie close to other QTLs (background QTLs) affecting the trait. In this case, each target locus is tested for association by multiple regressions in combination with a constant set of background loci (Manly and Olson, 1999). At each marker locus, the assessment of the strength of evidence for the

presence of a QTL is based on t-statistics or F-statistics. In a backcross, one may calculate a t-statistics to compare the averages of the two marker genotype groups. For other types of crosses (such as the intercross), where there are more than two possible genotypes, one uses a more general form of ANOVA, which provides F-statistics. The main advantages of ANOVA include its simplicity and there is no need for a genetic map for the markers because it considers each marker locus separately. However, the ANOVA approach for QTL mapping has four limitations (Lander and Botstein, 1989; Manly and Olson, 1999; Broman, 2001). First, it is difficult to conduct separate estimates of QTL location and QTL effect (proportion of phenotypic variance explained by the QTL). Second, individuals with missing genotypes often need to be discarded unless a mixed model that can handle unbalanced data and other statistical treatments is used. Third, when the markers are widely spaced and/or unevenly distributed, the QTL may be quite far from neighboring markers, and hence the power for QTL detection will decrease. Finally, there is a large amount of variation within each marker class and some of this will be due to other QTLs affecting the trait.

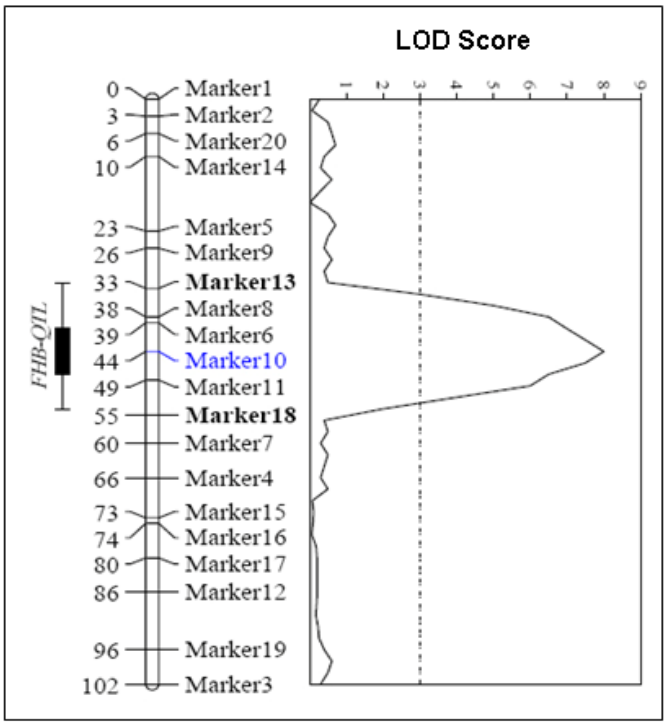

Fig. 7 The interval mapping approach for QTL mapping. The results of QTL mapping are plotted as a likelihood-ratio test statistic (LOD score) against the chromosomal map distance, measured in recombination units (centiMorgans). The vertical dotted line represents a threshold value above which a likelihood-ratio test provides a statistically significant fit to a model of the data. The best estimate of the location of the QTL is given by the chromosomal location that corresponds to the highest LOD score. In this hypothetical example, maximum LOD score is at 44 cM and the confidence interval is between 36 and 54 cM. Marker10 is the closest marker to the QTL while Marker13 and Marker18 are the two flanking markers.

Lander and Botstein (1989) developed a more powerful QTL mapping method, known as interval mapping (IM) that is often called simple interval mapping (SIM). Once a linkage map and phenotypic data are available for a population, SIM uses one marker-interval at a time to search for a hypothetical QTL (the target QTL) by performing a likelihood ratio test at every position within the interval. In this approach, the QTL is located within a chromosomal interval, defined by the flanking markers. Lander and Botstein (1989) proposed a simple rule for constructing confidence intervals for QTL position, which uses the likelihoods of odds (LOD score). LOD score is the base-10 logarithm of the ratio of two likelihoods (probabilities): the likelihood of the observed data assuming a QTL at the position in question and the likelihood assuming no QTL. The results of the analysis are plotted as a LOD score against the chromosomal map position in cM. The chromosomal location of the maximum LOD score is taken as the position of the QTL (Figure 7).

SIM has become the standard method used by many geneticists for mapping QTL and has been implemented in several freely distributed software packages (Basten et al. 2002 ; Manly and Olson, 1999). SIM procedure is based on maximum likelihood or regression and maximizes the likelihood of a single-gene genetic model by averaging over the possible states of the unknown genotype at each possible QTL location. SIM has more power and requires fewer progeny than ANOVA (Lander and Botstein, 1989; Haley and Knott, 1992; Zeng, 1994) but it has its own limitations. First, SIM considers one QTL at a time in the model (single-QTL model), ignoring the effects of other (mapped or not yet mapped) QTLs. Therefore, SIM can provide a biased identification and estimation of the effect and position of QTL when such multiple QTLs are located in the same linkage group (Haley and Knott, 1992; Knott and Haley, 1992; Martinez and Curnow, 1992; Zeng, 1994). Second, QTLs outside the interval under consideration can affect the ability to find a QTL within it (Zeng, 1993). Third, false identification of a QTL (false positive or 'ghost peak') can arise if other QTLs are linked to the interval of interest. Haley and Knott (1992) proposed a regression approach of interval mapping, which could save time in computation and produce similar results to those obtained by maximum likelihood but the estimate of the residual variance is biased and the power of QTL detection can be affected (Xu, 1995).

Multiple-QTL models are an improvement over single-QTL models because of their ability to separate linked QTLs on the same chromosome and to detect interacting QTLs that may otherwise be undetected (Schork, 1993). A variety of approaches have been proposed for mapping multiple QTLs. Jansen (1993), Zeng (1993) and Zeng (1994) independently proposed combining SIM with multiple regression analysis in mapping, which is termed as "composite interval mapping" (CIM). Like SIM, CIM evaluates the possibility of a target QTL at multiple analysis points across each inter-marker interval. However, at each point, it also includes the effect of one or more background markers that are often referred as cofactors. The purpose of using cofactors is to minimize the effects of QTLs in the remainder of the genome when attempting to identify a QTL in a particular region. The inclusion of cofactors in the analysis helps in one of two ways, depending on whether the background markers and the target interval are linked. If they are not linked, inclusion of the background markers 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 the target QTL from other linked QTL on the far side of the background marker (Zeng, 1993; Zeng, 1994). There are four major limitations in CIM: (i) CIM can be affected by an uneven distribution of markers in the genome (*i.e*., the test statistics in a marker-rich region may not be comparable to that in a marker-poor region); (ii) there is

difficulty of estimating the joint contribution to the genetic variance of multiple linked QTLs; (iii) CIM is not directly extendable for analyzing epistasis; (iv) the use of tightly linked markers as cofactors can reduce the statistical power to detect a QTL (Zeng et al. 1999).

To address the limitations of CIM, Kao et al. (1999) proposed and implemented multiple interval mapping (MIM) for mapping multiple QTLs simultaneously. The idea of MIM is to fit multiple putative QTL effects and associated epistatic effects directly in a model to facilitate the search, test and estimation of positions, effects and interactions of multiple QTLs. MIM consists of four components: (1) an evaluation procedure designed to analyze the likelihood of the data given a genetic model (number, positions and epistatic terms of QTL); (2) a search strategy optimized to select the best genetic model (among those sampled) in the parameter space; (3) an estimation procedure for all parameters of the genetic architecture of the quantitative traits (number, positions, effects and epistasis of QTL; genetic variances and covariances explained by QTL effects); and (4) a prediction procedure to estimate or predict the genotypic values of individuals and their offspring based on the selected genetic model and estimated genetic parameter values (Zeng et al. 1999). When compared with methods such as SIM and CIM, therefore, MIM tends to be more powerful and precise in detecting QTLs. The MIM model is based on Cockerham's model for interpreting genetic parameters and the method of maximum likelihood for estimating genetic parameters (Kao et al. 1999). Satagopan et al. (1996) and Sillanpää and Arjas (1998) used a Bayesian approach relying on a Markov chain Monte Carlo simulation to map multiple QTLs.

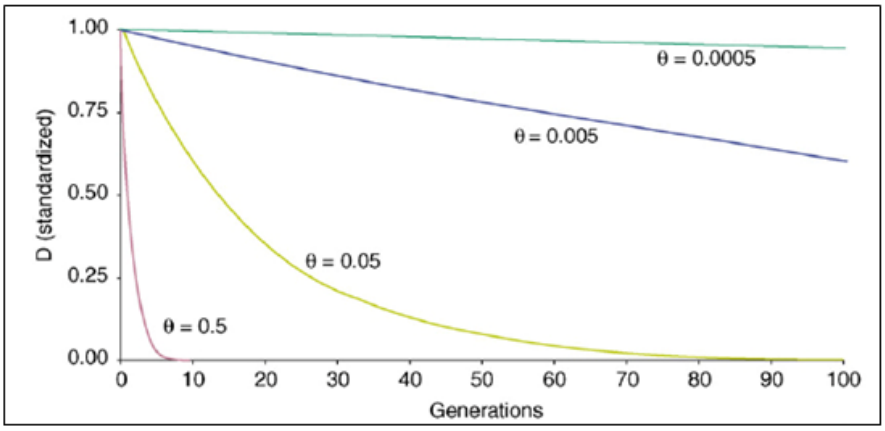

Fig. 8 Decay of linkage disequilibrium (LD) with time for four different recombination fractions (θ). For unlinked loci (θ = 0.5), LD decays rapidly within a small number of generations. For closely linked loci (θ < 0.05), the decay in LD is extremely slow. Abbreviation: *D* = coefficient of LD (Mackay and Powell, 2007).

All the different QTL mapping methods described above share a common assumption that the phenotype follows a normal distribution with equal variance in both parents. The least-squares and ANOVA based methods (Weller et al. 1990; Haley and Knott, 1992; Martinez and Curnow, 1992) assume that residual errors (*i.e.*, residuals within

QTL genotype classes) are normally distributed. Such methods are commonly said to be robust against non-normality. However, robustness against any type of non-normality in the context of QTL mapping methods has not been well established. On the other hand, maximum-likelihood based methods in interval mapping (Lander and Botstein, 1989) use the normal density function for the building of the likelihood itself. Quality of estimations is therefore very dependent on the normality of the phenotype. Many phenotypes of interest, however, are not normally distributed so the previously described QTL mapping methods cannot be directly applied in such cases. One approach to circumvent the assumption of normality is to use a mathematical transformation (*e.g.*, logarithm of 10, arcsine, etc.) that will convert the trait into an approximately normal distribution with equal variance in both parents (Wright, 1968). An alternative approach is to apply nonparametric (distribution free) statistical methods to QTL mapping that are applicable to any phenotypic distribution. Kruglyak and Lander (1995) described a non-parametric interval mapping approach based on the Wilcoxon rank-sum test applicable to experimental crosses. Coppieters et al. (1998) adapted this method to half-sib pedigrees in outbred populations.

Tests for QTL position and significance. One of the challenges for QTL mapping is the difficulty of determining appropriate significance thresholds (critical values) for the two types of errors: (a) that there is a segregating QTL whereas in reality there is not (false positive or type I error), and (b) that there is no QTL although it actually is present (false negative or type II error). The problem of determining appropriate threshold values appeared to be difficult because there are many factors that can vary from experiment to experiment and can influence the distribution of the test statistics. These include, but are not limited to, the sample size, the genome size of the organism under study, the genetic map density, segregation ratio distortions, the proportion and pattern of missing data, and the number and magnitude of segregating QTLs (Churchill and Doerge, 1994). Several papers addressed the problem of statistical significance in QTL analysis and presented solutions for hypothesis testing that are based on cumulative distribution functions of the LOD score (Lander and Botstein, 1989), permutation tests (Churchill and Doerge, 1994), bootstrap resampling method (Efron, 1979; Mammen, 1993) or a bootstrap model selection procedure (Shao, 1996).

Lander and Botstein (1989) used a “LOD drop-off method”, finding the location to each side of the estimated QTL location corresponding to a decrease of one from the maximum LOD score. Figure 7 illustrates this using a hypothetical data. The maximum LOD score is at 44 cM, and the confidence interval is between 36 and 54 cM. Most researchers use a minimum LOD score of 3.0 or its equivalent to declare a QTL. The introduction of different resampling methods, such as permutation tests (Churchill and Doerge, 1994; Doerge and Churchill, 1996), bootstrap resampling method (Efron, 1979; Mammen, 1993; Visscher et al. 1996), bootstrap model selection procedure (Shao, 1996) and cross validation (Utz et al. 2000) provided a computationally simple and free of dubious assumptions for establishing the significance threshold value. Permutation tests generate many different samples from the actual data by "shuffling" the trait values with respect to the marker genotypes to estimate empirically the threshold for a test statistic for detection of a QTL. This approach accounts for missing marker data, actual marker densities, and nonrandom segregation of marker alleles. A permuted sample is generated from the data by randomly pairing phenotypes and genotypes in the sample, stimulating the null hypothesis of no intrinsic association between genotypes and phenotypes (no QTL). The statistical test is then performed over the whole genome on the permuted sample for QTL, and the maximum test statistics is recorded. This permutation analysis is

repeated for a number of replicates (usually 1,000 permutations) to obtain a distribution of the maximum test statistics, and from the distribution to obtain the threshold value. One then compares this threshold with the test statistics from the original sample, and declares the existence of a QTL if the peak test statistics in a region exceeds the threshold. Subsequently, Doerge and Churchill (1996) extended the permutation method for detecting multiple QTLs using conditional empirical threshold or residual empirical threshold. Compared with the standard permutation test, the latter two methods tend to have greater statistical power. However, the methods are not designed for detecting multiple linked QTLs.

For MIM, where model selection is involved, Zeng et al. (1999) proposed using a bootstrap re-sampling method for hypothesis testing. However, the heavy computational burden has limited the use of the bootstrap test. Furthermore, it is unclear how to apply the bootstrap method in Zeng et al. (1999) to the situation where a nonlinear model, such as logistic regression or a Poisson model, is used to map multiple QTLs with MIM, since the bootstrap procedure is performed on model-based residuals. The importance of cross-validation (CV) has been strongly emphasized by Utz et al. (2000). CV is a technique for assessing how the results of a statistical analysis will generalize to an independent dataset. It is done by partitioning the data into complementary subsets for performing the initial analysis and validation set for validating the analysis. Utz et al. (2000) showed that the proportion of genotypic variance explained in QTL-models based on composite interval mapping, may be over-optimistic. All such analyses should include CV against other environments, re-sampled genotypes from the same population or both.

Limitations of linkage-based QTL mapping methods. The linkage analysis-based QTL mapping methods have the following limitations. First, the need for evenly distributed marker spacing of 10-20 cM (Darvasi et al. 1993) and a high number of informative individuals may make linkage a somewhat limited and even unsuccessful procedure (Lander and Kruglyak, 1995). It should, however be noted that dense markers may pose problems for linkage analysis software's in providing correct marker order and can lead to erroneous QTL mapping results as has recently been reported by Collard et al. (2009). Second, the parents used to develop the mapping population may be out-of-date by the time when the genotype and phenotype data is available. Many marker development projects for annual crops are using populations that were established five or more years before the genotyping work. This could reduce the value of the information gathered and the scope of its implementation. Third, QTL identification based on linkage studies identify chromosomal regions, not individual genes, which may affect a trait. Linkage analysis in plants typically localizes QTLs within 10 to 20 cM intervals because of the limited number of recombination events that occur during the construction of mapping populations and the cost for propagating and evaluating a large number of lines (Doerge, 2002; Holland, 2007). For species with large genome size, this large interval may contain many genes unless the chromosomal region associated with the trait is fine mapped by genotyping large population size with thousands of high throughput and low cost markers, such as single nucleotide polymorphic (SNP) markers.

Association mapping

Linkage disequilibrium-based association analysis is the second QTL mapping method. The terms linkage disequilibrium (LD) and association mapping have often been used interchangeably in literature. According to Gupta et al. (2005), however,

association mapping refers to significant association of a molecular marker with a phenotypic trait while LD refers to non-random association between two markers (alleles at different loci), between two genes or QTLs, between a gene/QTL and a marker locus. Thus, association mapping is one of the several uses of LD. The terms linkage and LD are also often confused. Linkage refers to the correlated inheritance of loci through the physical connection on a chromosome, whereas LD refers to the correlation between alleles in a population (Flint-Garcia et al. 2003) but not necessarily on the same chromosome. LD can be used in plant genomes for construction of LD maps, for studying marker-trait association both independently and in combination with linkage analysis, and for the study of population genetics and evolution both in nature and under domestication (Gupta et al. 2005).

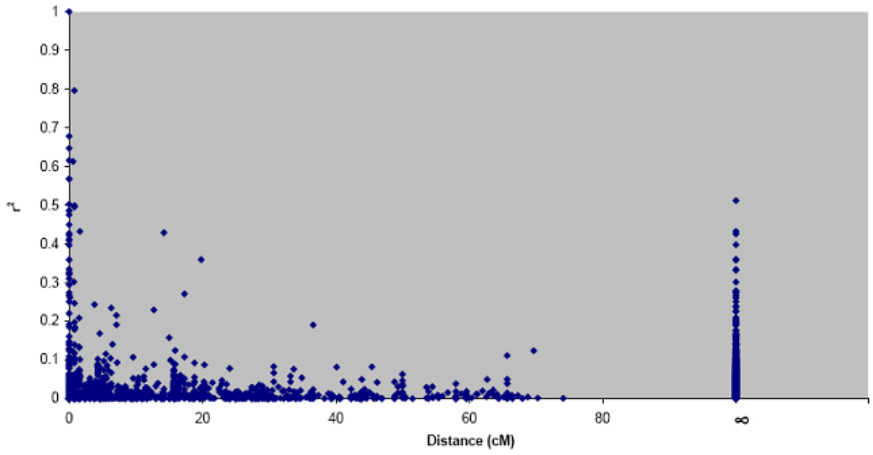

Figure 9. Linkage disequilibrium (LD) decay plot. The squared correlation between paired marker intensities on the y-axis is plotted against the distance between pairs of markers in centi-Morgan (cM) on the x-axis. LD between pairs of markers on different non-homologous chromosomes is represented by the symbol ∞. The commonly accepted reference value for r^2 of 0.1 indicates LD to decay at about 3 cM. The critical test value for r^2 is 0.03 for a test level of 0.01. Using this critical value as threshold, LD decays at about 8 cM (D'Hoop et al. 2008).

Association mapping generally falls into two broad categories: (i) candidate-gene association mapping, which relates polymorphisms in selected candidate genes that have purported roles in controlling phenotypic variation for specific traits; and (ii) genome-wide association mapping, or genome scan, which surveys genetic variation in the whole genome to find signals of association for various complex traits (Risch and Merikangas, 1996). For candidate-gene association mapping, information regarding the location and function of genes involved in genetic, biochemical or physiological pathways that lead to final trait variation is often required (Risch and Merikangas, 1996; Mackay, 2001). Candidate-gene association mapping requires the identification of SNPs between lines and within specific genes because SNPs offer the highest resolution for mapping QTLs and are potentially in LD with the causative polymorphism (Rafalski, 2002). Whole-genome association scans requires high-capacity DNA sequencing instruments or high-density oligonucleotide (oligo) arrays to efficiently identify SNPs at a density that accurately reflects genome-wide LD structure and haplotype diversity. As sequencing and genotyping costs continue to

decrease, we expect to see more genome-wide association mapping. As reviewed by Zhu et al. (2008), population size for several association mapping studies is about 100, which is much lower compared to individuals used for linkage-based QTL mapping. However, Zhu et al. (2008) simulations with empirical maize data show that a large sample size is required to obtain high power to detect genetic effects of moderate size.

Association mapping is a population-based survey that capitalizes on historical recombination to identify candidate genes affecting complex traits (Falconer and Mackay, 1996). Unlike linkage analysis, where familial relationships are used to predict correlations between phenotype and genotype, association mapping rely on previous, unrecorded sources of disequilibrium to create population-wide marker-phenotype associations (Kruglyak, 1999; Ewens and Spielman, 2001; Jannink et al. 2001). Genetic diversity is evaluated across natural populations to identify polymorphisms that correlate with phenotypic variation. LD is seen in large populations over many generations when selective pressure increases or decreases the frequency of particular alleles or allelic combinations (Falconer and Mackay, 1996) (Figure 8). Association mapping and linkage mapping differ in terms of how the genetic architecture of the trait affects statistical power. The importance of a particular allele in phenotypic variation across a population depends on its frequency, as well as on its effect. Thus, association mapping has low power to detect rare alleles, even if these alleles have a large phenotypic effect. Conversely, alleles that are identified by linkage-based mapping that involves crosses between two essentially randomly chosen parents can have a large effect but might not be important from an evolutionary perspective because they are rare. Thus, mapping QTLs by using crosses might bias researchers towards identifying rare (and often perhaps deleterious) alleles that have large effects but little relevance to most of the phenotypic diversity found in nature (Nordborg and Weigel, 2008). Linkage analysis with experimental populations derived from a bi-parental cross provides pertinent information about traits that tends to be specific to the same or genetically related populations, while results from association mapping are more applicable to a much wider germplasm base. The ability to map QTLs in collections of breeding lines, landraces, or samples from natural populations has great potential for future trait improvement and food security.

However, several factors contribute to false positives and affect the success of LD mapping: (i) there are a number of factors that can lead to an increase in LD (e.g., population structure or subdivision, population admixture, population bottleneck or small population size, natural and artificial selection, inbreeding, genetic isolation between lineages, and low recombination rate); (ii) some factors lead to a decrease or disruption in LD (*e.g.*, outcrossing, high recombination rate, and high mutation rate); and (iii) other factors may lead to either an increase or a decrease in LD, or may increase LD between some pairs of alleles and decrease LD between other pairs (*e.g.*, mutations and genomic rearrangements). Several statistical methods have been proposed to account for population structure and familial relatedness, structured association (Pritchard and Rosenberg, 1999; Pritchard et al. 2000; Falush et al. 2003), genomic control (Devlin and Roeder, 1999), mixed model approach (Yu et al. 2006), and principal component approach (Price et al. 2006).

Table 2. **The most commonly used QTL mapping statistical programs for both linkage-based and linkage disequlibrium-based mapping methods, their operating system, and online links.**

Name	Version	Platform (operating system)	Description	Availability	References	Online links (verified in July 2009)
Map Manager QTX	b29	Windows, Mac OS	A graphic, interactive program to map quantitative trait loci using intercrosses, backcrosses or recombinant inbred strains in experimental plants or animals.	Free	Manly and Olson, 1999.	http://www.mapmanager.org/mmQTX.html
Mapmaker/ QTL	1.1	UNIX, VMS, DOS, Mac OS	A package containing a program for genetic linkage analysis and a program for mapping genes underlying complex traits.	Free	Lincoln et al. 1992.	http://www.broadinstitute.org/ftp/distribution/software/mapmaker3/
MapQTL	5	Windows ® (95/98/ME/NT4.0/2000/XP/Vista 32-bit	Mapping of quantitative trait loci (QTL) for several types of experimental mapping populations.	Comercial	Van Ooijen, 2009.	http://www.kyazma.nl/index.php/mc.MapQTL/sc
PlabQTL	1.2	DOS	A program characterizing loci that affect the variation of quantitative traits.	Free	Utz and Melchinger, 2003.	https://www.uni-hohenheim.de/plantbreeding/software/
QGene	4.0	Any computer	An entirely rebuilt Java application that will run on any computer.	Free	Nelson, 1997.	http://www.qgene.org/
QTL Cartographer	2.5 for Windows	UNIX, DOS, Windows, Mac OS	A program to map quantitative traits using a map of molecular markers.	Free	Basten et al. 1994; Wang et al. 2007.	http://statgen.ncsu.edu/qtlcart/index.php
Structure	2.3	DOS, Windows, UNIX (Solaris), Linux	A program for investigating population structure.	Free	Pritchard et al. 2000.	http://pritch.bsd.uchicago.edu/structure.html
TASSEL	2	Web-based (operating system independent)	A software package for association mapping, diversity estimation and calculating linkage disequilibrium.	Free	Zhang et al. 2006.	http://sourceforge.net/projects/tassel

LD can be caused by admixture of subpopulation, which leads to false-positive results if not correctly controlled in statistical analysis. Such false-positives arise when testing random genetic markers with different frequencies in subpopulations for a trait with parallel phenotypic differences. If, however, the distribution of functional alleles is highly correlated with population structure, statistically controlling for population structure can result in false-negatives, particularly for small sample sizes. Although population subdivision increases LD, its effect depends on the number of populations, the rate of exchange between populations, and the recombination rate (Pritchard and Przeworski, 2001). Association studies, therefore, are best carried out in independent populations with a large sample size (Yu and Buckler, 2006). Admixture is gene flow between individuals of genetically distinct populations followed by intermating. Admixture results in the introduction of chromosomes of different ancestry and allele frequencies. Often, the resulting LD extends to unlinked sites, even on different chromosomes, but breaks down rapidly with random mating. In small populations, the effects of genetic drift result in the consistent loss of rare allelic combinations, which increase LD levels but in the absence of other mitigating factors (such as population subdivision), this effect should be short-lived (Wall et al. 2002). The same is true of directional selection; strong selection for a particular allele limits genetic diversity around a locus, resulting in a short-term increase in LD around the selected gene. Effective detectable recombination rate and LD decay decreases as homozygosity increases. Thus, LD decays more rapidly in outcrossing species as compared to selfing species (Nordborg, 2000) because recombination is less effective in selfing species, where individuals are more likely to be homozygous, than in outcrossing species. Mutations will disrupt LD between pairs involving wild alleles, and will promote LD between pairs involving mutant alleles. Similarly, genomic rearrangements may disrupt LD between genes separated due to rearrangement, but LD may increase between new gene combinations in the vicinity of breakpoints due to suppression of local recombination. All of the various factors described above affect the utility of LD for localizing QTLs. Any evolutionary force that increases LD beyond that expected by chance in an ideal population will inflate the rate of false-positive associations.

Statistical measures for LD. The basic statistics for association analysis, under an ideal situation, would be linear regression, analysis of variance (ANOVA), t-test or chi-square test. However, as population structure can generate spurious genotype-phenotype associations, different statistical approaches have been designed to deal with this confounding factor. The different measures and methods for estimating the level of LD includes two-locus methods, transmission disequilibrium test, admixture disequilibrium mapping, least-squares multilocus method, and haplotype segment sharing methods (Jorde, 2000). These methods have been described in different reviews (Jorde, 2000; Flint-Garcia et al. 2003; Gaut and Long, 2003; Gupta et al. 2005; Mackay and Powell, 2007).

Although a variety of statistics have been used to measure LD, the two most commonly used statistics are r^2 (square of the correlation coefficient) and D' (disequilibrium coefficient). The statistics r^2 and D' reflect different aspects of LD and perform differently under various conditions. Whereas r^2 summarizes both recombinational and mutational history, D' measures only recombinational history and is therefore the more accurate statistics for estimating recombination differences. However, D' is strongly affected by small sample sizes, resulting in highly erratic behavior when comparing loci with low allele frequencies. For the purpose of examining the resolution of association studies, Flint-Garcia et al. (2003) suggested using r^2 statistics, as it is indicative of how markers might correlate with the QTL of

interest. The two common ways to visualize the extent of LD between pairs of loci are (a) LD decay plots (Figure 9) and (b) disequilibrium matrices. LD decay plots are used to visualize the rate at which LD declines with genetic or physical distance (Figure 9). Scatter plots of r^2 values versus genetic/physical distances between all pairs of alleles within a gene, along a chromosome, or across the genome are constructed. The decay of LD over physical distance in a population determines the density of marker coverage needed to perform an association analysis. Figure 10 illustrates the number of markers required for association mapping in 3 crops of different genome size (rice = 466 x 10^3 kilobase pair, abbreviated as kb; Sorghum = 735 x 10^3 kb; maize = 2400 x 10^3 kb). If LD decays rapidly, then a higher marker density is required to capture markers located close enough to functional sites (Flint-Garcia et al. 2003; Gaut and Long, 2003).

The extent and patterns of linkage disequilibrium have been characterized in several crop species. In maize, r^2 decays within 0.3-2 kb, and this rapid decay may be due to its outcrossing mating system (Remington et al. 2001; Tenaillon et al. 2001). In commercial maize inbred lines, LD decay may be slower and linkage blocks may extend more than 100 kb (Ching et al. 2002). For regions that have experienced strong selective sweeps, LD may extend over 500 kb (Jung et al. 2004). Yan et al. (2009) genotyped 632 lines with 1229 SNPs that represent 538 loci and reported an average LD decay ranging from 5 kb to 10 kb depending on the chromosome. In sorghum, $r^2 > 0.1$ is observed between 15 to 20 kb (Hamblin et al. 2005). High levels of marker association ($r^2 > 0.1$) across a 212 kb region was observed in cultivated, elite varieties of barley, while in landrace accessions, LD levels persist to about 90 kb (Caldwell et al. 2006). In rice, LD decay of about 100 kb has been reported around a disease resistance locus (Garris et al. 2003). Using unlinked SNPs to determine the amount of background linkage disequilibrium in different rice population, Mather et al. (2007) reported LD decay of about 500 kb in temperate japonica, 150 kb in tropical japonica and 75 kb in indica. These results, together with others, clearly demonstrated the high variability in LD decay depending upon the species, populations, or genomic region under consideration. A single study determining LD for a species can therefore not be projected to all populations of the species. Inference of LD levels across the genome of a population can also be misleading because LD patterns are variable among chromosomes and over distance.

Advantages and disadvantages of LD. There are six potential advantages of the LD-based QTL mapping in crop species. First, it increases mapping resolution, reduced research time, and reveal greater allele number (Yu and Buckler, 2006). Second, LD mapping provides detailed marker data points on a large number of lines and varieties, which will be valuable in several breeding strategies. Third, the LD approach uses real breeding populations, the material is diverse and relevant, and the most important genes should be segregating in such populations (Figure 5). Plant breeders are often reluctant to grow and assess a huge number of lines with little or no potential for direct commercial outcome such as required for genetic map construction and fine mapping of QTLs. The advantage of LD mapping to the breeder is that mapping and commercial variety development is conducted simultaneously. Fourth, pattern analysis of marker data might detect complex combinations (even epistatic interactions) between alleles at several loci, which underlie the superior individuals in a breeding population. This might prove difficult to isolate and validate using the linkage-based mapping approach. Fifth, LD studies have proven to be more powerful for genes of small to modest effects (Risch and Merikangas, 1996), reduce sample requirements in terms of size and structure as compared to those needed for linkage studies to obtain similar significance, and narrow the distance between the

expected site of the locus and the nearest marker. Finally, LD approach has the potential not only to identify and map QTLs (Meuwissen and Goddard, 2000), but also to identify (a) causal polymorphism within a gene that is responsible for the difference in two alternative phenotypes (Palaisa et al. 2003, Palaisa et al. 2004), and (b) haplotype blocks and haplotypes representing different alleles of a gene and their combinations.

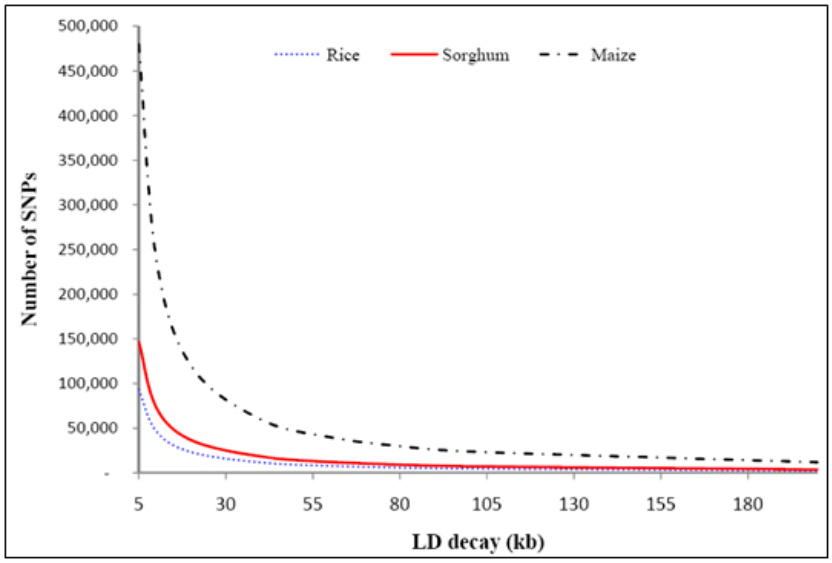

Figure 10. Number of SNP markers required for association mapping at different LD decays for 3 crops of different genome size (rice: 466 megabases, abbreviated as Mb; sorghum: 735 Mb; Maize: 2400 Mb).

However, LD-based mapping has several potential disadvantages. First, it assumes that the trait of interest is segregating in the breeding material and hence may not assist in the identification and introgression of novel alleles. Therefore, there will be a continuing requirement for advanced backcross QTL (AB-QTL) mapping for introgression of novel alleles from wild relatives (Tanksley and Nelson, 1996) and a capability for map construction for other special cases. Second, LD mapping strategies will work best where there is strong selection pressure for the trait of interest, so the location and management of field trials and the design and application of laboratory assays is crucial to its success. Third, LD mapping provides little insight into the mechanistic basis of LD detected (*e.g.*, LD may not be due to linkage in all cases) so that genomic localization and cloning of genes based on LD may not be successful, particularly for those with relatively small effect. This is because a strong LD may sometimes be due to recent occurrence of LD rather than a close physical linkage between the two loci. Fourth, several factors can affect the reliability of LD mapping in plants, including population structure or subdivision, population admixture, population bottleneck (small population size), natural and artificial selection, inbreeding, genetic isolation between lineages, recombination rate, outcrossing, mutation, and genomic rearrangements (*e.g.*, Gaut and Long, 2003; Gupta et al.

2005). Fifth, LD mapping often requires genotyping the mapping populations with large number of markers (Yan et al. 2009). The number of markers required for whole-genome scan depends on the genome size of the species and the expected LD decay. If LD decays at 5 kb, for example, the number of SNPs required for genotyping will be as high as 93200 for rice, 147000 for sorghum, 480000 for maize, 1.1 million for barley and 3.2 million for hexaploid wheat. The number of markers will decrease sharply if LD decay is expected at about 100 kb (4660 SNPs for rice, 7350 SNPs for sorghum, 24000 SNPs for maize, 57000 SNPs for barley and 160000 SNPs for hexaploid wheat). Such high density marker requirement can only be achieved through the development of an integrated genotyping by sequencing platform and analytical tools.

Other QTL mapping methods

Both the linkage-based and LD-based QTL mapping methods have their own limitations when used alone. A new joint linkage and LD mapping strategy has been devised for genetic mapping, taking advantage of each approach (Wu and Zeng, 2001; Wu et al. 2002). The approach of combined linkage analysis and LD for QTL analysis has been extended for multi-trait fine mapping of QTLs (Lund et al. 2003; Meuwissen and Goddard, 2004; Gupta et al. 2005).

Multivariate analysis for multi-trait QTL detection in inbred lines has been proposed by different authors (Jiang and Zeng, 1995; Korol et al. 1995; Weller et al. 1996; Knott and Haley, 2000). Ronin et al. (1995) have extended this to half sib families. Weller et al. (1996) proposed to synthesize most of the information in a linear combination of the traits and used principal component analysis (PCA) for multi-trait detection of QTLs in dairy cow. PCA was then used for mapping QTLs in genetic crosses (Liu et al. 1996; Zeng et al. 2000; Gilbert and LeRoy, 2003; Upadyayula et al. 2006) and association analysis in pedigrees (Chase et al. 2002). Yan et al. (2003) have developed a PCA-based program called ‘‘GGE analyses for the genetic analysis of GxE interactions, QTL and diallels in barley". Bjørnstad et al. (2004) described the utility of bi-linear modeling by cross-validated partial least squares regression (PLSR) for exploring the relationship between genotype and phenotype.

Statistical programs for QTL mapping

A large quantity of mapping data can now be produced at an unprecedented rate, requiring the use of dedicated computer programs to extract all embedded information. Several statistical packages have been developed for QTL mapping in the last two decades, which are among about 400 genetic analysis software’s that are listed at http://www.nslij-genetics.org/soft/ and offer a panel of standard and more sophisticated analyses. The review below provides a brief outline of some of the most commonly used statistical software’s for QTL mapping, including their operating systems and their online links (Table 2).

Statistical programs for linkage-based mapping

1. QTL Cartographer (Basten et al. 1994; Basten et al. 2002; Wang et al. 2007) is a suite of programs for DOS, UNIX, MacOS or Windows. QTL Cartographer is distinguished by its menu-driven interface, its more detailed documentation, and its resampling methods. Windows QTL Cartographer

(WinQTLCart) maps QTLs in cross populations from inbred lines. WinQTLCart includes a graphic tool for presenting mapping results and can import and export data in a variety of formats. This program implements the following statistical methods: single-marker analysis, interval mapping, composite interval mapping, Bayesian interval mapping, multiple interval mapping, multiple trait analysis, and multiple trait MIM analysis.

2. MAPMAKER/QTL (Lincoln et al. 1992) is a widely used program for UNIX, DOS and Mac OS operating systems. Researchers need to follow three basic stages for doing QTL analysis with Mapmaker/QTL: (i) prepare the data into the format that Mapmaker needs; (ii) constructing genetic map for the marker data (establish linkage groups, calculate map distances and determine locus orders) with MAPMAKER/EXP (Lincoln et al. 1992), and (iii) feed the marker and phenotype data along with marker distances for each linkage group into Mapmaker/QTL and run QTL analyses.
3. PlabQTL (Utz and Melchinger, 2003) is a script-driven program for DOS or AIX that is designed to analyze automatically a dataset at increasing levels of complexity in successive runs. Like MAPMAKE/QTL, researchers need to construct genetic map for the marker data, prepare a complete input file that consists of the marker data, the linkage map, and the phenotypic observation values. The complete data set needs to be accompanied by an analysis controlling command file that should instruct the program to perform either simple interval mapping or composite interval mapping.
4. QGene (Nelson, 1997), version 4.0, is a program with a variety of graphics for displaying analyses outputs. These functions make it uniquely useful for rapid exploration of data using any computer. However, it does not perform CIM.
5. Map Manager QTX (Manly et al. 2001) is a Mac OS or Windows based program for the analysis of genetic mapping experiments in experimental plants and animals. It includes functions for mapping both Mendelian and quantitative trait loci. QTX is an enhanced version of Map Manager QT, which was designed to be used either as a mapping program itself or as a data-preparation program for other mapping programs.
6. MapQTL (Van Ooijen, 2009) is a Windows-based commercial program that is distinguished by its ability to map QTLs in populations derived from both inbred parents (BC_1, F_2, RIL, DH) and non-inbred parents (full-sib family) in which both markers and QTL may have more than two alleles. It also offers a nonparametric form of single-locus association, the Kruskal-Wallis rank sum test, which is applicable for data with distributions far from normal. The input data for MapQTL is a plain text files with a flexible layout of the quantitative trait data, the molecular marker genotypes and the linkage map.
7. Epistat (Chase et al. 1997) is a DOS operating interactive program that combines statistical methods and color-graphic. The program organizes genetic mapping data and quantitative trait values into color graphic displays which illustrate the individual effects of a single QTL as well as the interactions between pairs of QTLs. For a given trait, the program displays the effects of the alleles at each of two loci on the quantitative trait value, as well as the effects of the interactions between these alleles. Log likelihood ratios are used to compare the likelihood of explaining the effects by null, additive, or epistatic models. This program does not perform interval mapping and therefore does not require prior genetic map construction.

Statistical programs for LD-based mapping. Trait Analysis by aSSociation, Evolution, and Linkage (acronym - TASSEL), (Zhang et al. 2006) is the most

commonly used software for association mapping in plants and is frequently updated as new methods are developed (Bradbury et al. 2007). In addition to association analysis (*i.e*., logistic regression, linear model, and mixed model), TASSEL is also used for calculation and graphical display of linkage disequilibrium statistics, browsing and importation of genotypic and phenotypic data, and evolutionary patterns. The software has three advantages: (i) it integrates various diversity databases, including Panzea (http://www.panzea.org), Gramene (http://www.gramene.org), and Germinate (http://bioinf.scri.ac.uk/germinate/wordpress), by means of Genomic Diversity and Phenotype Connection (GDPC) middleware; (ii) it provides new statistical approaches to association mapping such as a General Linear Model (GLM) and Mixed Linear Model (MLM), and (iii) it has the ability to handle a wide range of indels (insertion & deletions), which is ignored by most software package.

SAS software (SAS Institute, 1999) and R (Ihaka and Gentleman, 1996) often are used by advanced researchers with programming skills as the platform to develop various methods. SAS/GENETICS is a Windows, UNIX(HPUX/AIXR/Solaris) and MVS based commercial program added to the SAS System software for summarizing marker properties (allele and genotype frequencies, tests for Hardy-Weinberg equilibrium, measures of marker informativeness), examining marker-marker relationships (tests and measures of linkage disequilibrium, and haplotype frequency estimation), and exploring marker-trait associations using case-control or family-based tests. The complex breeding history of many important crops and the limited gene flow in most wild plants has created population structure or stratification within the germplasm (Sharbel et al. 2000). Since population structure tends to create spurious LD between unlinked markers (Nei and Li, 1973; Pritchard and Przeworski, 2001), scientists need to conduct complementary analyses on genotypic data to assess for population structure before proceeding with LD analysis. Various statistical programs have been described in the literature for assessing population structure (*e.g*., Jorde, 2000; Gupta et al. 2005) but only some of them have been used in plants. These include STRUCTURE (Pritchard et al. 2000), Arlequin (Excoffier et al. 2005), FSTAT (Goudet, 1995), GENEPOP (Raymond and Rousset, 1995), Genetic Data Analysis (GDA), (Weir, 1996), and GENETIX (Belkhir et al. 1996-2004). For example, STRUCTURE is a DOS, Windows, UNIX (Solaris) and Linux based program that implements a model-based clustering method for inferring the presence of population structure, identifying distinct genetic populations, assigning individuals to populations, and identifying migrants and admixed individuals.

PROGRESSES AND FUTURE PROSPECTS

Since the early 1990s, numerous studies have identified molecular markers linked to QTLs involved in the inheritance of agronomically important traits in a wide range of crop species. Following the discovery of promising QTLs and identification of molecular markers, MAS has been used to transfer single genes or QTL in various species. However, published results in QTL introgressions through MAS are variable, ranging from successful experiments to those with limited success and even a failure (see Semagn et al. 2006b for review). The rate of success starts to decrease when five or more target QTLs for complex traits are introgressed in to a given germplasm (Lawson et al. 1997; Shen et al. 2001; Bouchez et al. 2002; Ribaut et al. 2002a; Lecomte et al. 2004; Thabuis et al. 2004). Several factors may contribute for such failure or unexpected results in MAS: (i) errors in QTL mapping (the putative QTL may be a false positive or the QTL effect might have been over estimated); (ii) the

repeatability of the QTL across different genetic background and/or environments might have not be confirmed (*e.g*., Melchinger et al. 1998; Schon et al. 2004); (iii) there may be QTL by environment and QTL by QTL interactions (*e.g*., Ribaut et al. 2002a, Ribaut et al. 2002b); (iv) pleiotrophic effects (Tuberosa et al. 2002); and (v) the chromosomal segments associated with QTL hold not just one but several genes, and recombination between those genes would then modify the effect of the introgressed segments (*e.g*., Eshed and Zamir, 1995; Monna et al. 2002). For example, Kroymann and Mitchell-Olds (2005) find mapped phenotypic effects segregating within a one-centimorgan chromosome interval in *Arabidopsis thaliana* for which lines with mapped recombination breakpoints were available, and examined the sequence signature of historical polymorphism. The authors found that the 1 cM chromosome interval contained two growth rate QTLs within 210 kilobases (kb). Both QTLs showed epistasis (*i.e*., their phenotypic effects depended on the genetic background). This amount of complexity in such a small area suggests a highly polygenic architecture of quantitative variation, much more than previously documented (Koornneef et al. 2004).

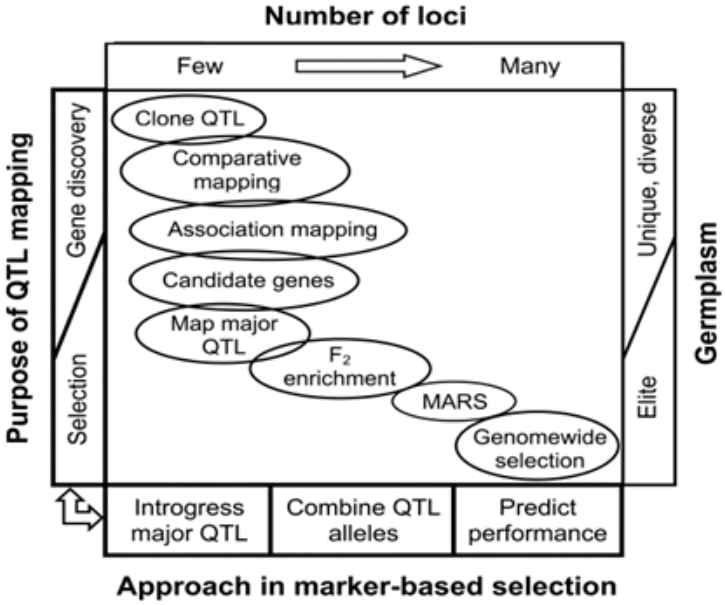

Figure 11. Goals and approaches for using molecular markers to study and select for complex traits in plants (Bernardo, 2008).

Overestimation of the effect of QTLs is often a major problem because QTLs whose effects are overestimated are more likely to be detected above the necessary stringent threshold than are those whose effects are correctly estimated or underestimated. Spurious QTLs (false positives) are occasionally detected and this represents an overestimation of an effect that is actually zero (Haley and Andersson, 1997). Furthermore, most primary or coarse QTL mapping studies using small population size and low marker density allows only for an approximate mapping of the chromosomal region. Therefore, identification of reliable QTL is a preliminary step in

developing a marker assisted selection programs for genetic improvement. To utilize QTLs in selective breeding or to identify functional genes, the identified major QTLs should be fine mapped to a higher level of resolution for QTL position and verified or validated in additional genetic backgrounds and environments by developing advanced segregating populations with large numbers of recombinations in the region of interest. QTL verification is defined as the repeated detection of the same marker alleles at a similar position on the genetic map of a chromosome, of a QTL controlling a trait under more than one set of experimental conditions (Brown et al. 2003). Verification of QTL is necessary to substantiate a biological basis for observed marker-trait associations, to provide precise estimates of the magnitude of QTL effects, and to predict QTL expression at a given age or in a particular environment. Only then will sufficient experimental evidence be in place to monitor the transmission of trait genes via closely linked markers as a selection criterion (Young, 1999).

Fine mapping of major QTLs requires the construction of special populations with large numbers of recombinations in the region identified by the coarse genome scan (Darvasi, 1998; Pumphrey et al. 2007). A widely adopted strategy to estimate the position and effect of a coarsely mapped QTLs more accurately is to create a new experimental population by crossing nearly isogenic lines (NILs) that differ only in the allelic constitution at the short chromosome segment harbouring the QTL (QTL-NILs). In such a population, because of the absence of other segregating QTLs, the target QTL becomes the major genetic source of variation, and the phenotypic means of the QTL genotypic classes (+/+, -/- and, when present, +/-) can be statistically differentiated and genotypes recognized accordingly (Salvi and Tuberosa, 2005).

However, there are at least three shortcomings regarding QTL fine mapping and validation. First, time and effort required for developing NILs, introgression libraries, advanced backcross QTL (AB-QTL) introgression lines, as well as the limited genetic variability as a result of using only two parental lines are crucial aspects to be considered. Second, the time and cost required to genotype, and adequately phenotype the fine mapping and validation populations also represents a substantial investment and slows the application of marker information to genetic improvement (Pumphrey et al. 2007). Third, it is almost impossible to fine map several minor QTLs associated with highly complex traits, such as drought tolerance and yield for different reasons: (a) the magnitude of inconsistency in estimated QTL effects is much higher for complex traits controlled by many minor QTLs rather than by a few major QTLs; (b) most QTLs are often background or germplasm specific and estimated QTL effects will have limited transferability across populations (*i.e*., QTL mapping for such traits will likely have to be repeated for each breeding population); (c) as complex traits controlled by many QTLs are subject to genotype -by- environment interaction, QTL mapping for the same population will likely have to be performed for each target set of environments; and (d) as the effects of sampling error are large, population sizes of 500-1000 are needed for mapping QTLs *per se* (Bernardo, 2008). In such cases, it is unlikely that a “QTL hunt”, involving traditional mapping of QTLs in a small number of crosses, with the objective of tracking useful alleles that will have consistent and large effects in other backgrounds via MAS, will be successful.

To overcome some of the problems in fine mapping, and/or QTL validation, marker assisted recurrent selection (MARS) (Edwards and Johnson, 1994; Hospital et al. 1997; Johnson, 2004; Bernardo and Charcosset, 2006) and genomewide selection (also called genomic selection), (Meuwissen et al. 2001; Bernardo and Yu, 2007; Wong and Bernardo, 2008; Bernardo, 2009; Heffner et al. 2009; Zhong et al. 2009;

Jannink et al. 2010) have been proposed for accumulating favorable alleles from many QTLs (up to 100 QTLs based on simulation studies) for highly polygenic traits (Bernardo and Yu, 2007; Bernardo, 2008). MARS refers to the improvement of an F_2 population by one cycle of marker-assisted selection (*i.e*., based on phenotypic data and marker scores) followed by three cycles of selection based on marker scores only (Johnson, 2001; Johnson, 2004). The marker scores are typically determined from about 20 to 35 markers that have been identified, in a multiple-regression model, as significantly associated with one or more traits of interest (Koebner, 2003). Genomewide selection refers to marker-based selection without significance testing and without identifying of a subset of markers associated with the trait (Meuwissen et al. 2001). It focuses purely on prediction of performance and avoids QTL mapping altogether (Bernardo, 2008; Heffner et al. 2009).

The current trend in molecular breeding is to combine the different QTL mapping methods (Figure 11) with methods in functional genomics (Varshney et al. 2006) and QTL cloning (Tuberosa and Salvi, 2006). The genetic maps of many organisms are now becoming increasingly dense, and the cost of genotyping is decreasing. The development of high resolution maps facilitates the isolation of actual genes or quantitative trait nucleotide (QTN) (rather than markers) via map-based cloning (also called positional cloning). The identification of genes controlling important traits will enable plant scientists to predict gene function using reverse genetics methods (*e.g*., TILLING, Eco-TILLING; (McCallum et al*.* 2000; Comai et al. 2004), isolate homologues and conduct transgenic experiments. The use of gene sequences derived from expressed sequence tags (ESTs) or gene analogues, described as the 'candidate gene approach', holds promise in identifying the actual genes that control the desired traits (Yamamoto and Sasaki, 1997; Cato et al. 2001; Pflieger et al. 2001). The number of EST and genomic sequences available in databases is growing rapidly (especially from genome sequencing projects), and the accumulation of these sequences will be extremely useful for the discovery of single nucleotide polymorphisms (SNPs) and data mining for new markers in the future (Gupta et al. 2001; Kantety et al. 2002).

To enhance the efficiency of MAS, knowledge of the DNA sequence of the gene enables the design of direct markers, which are located within the actual gene, thus eliminating the possibility of recombination between marker and gene (Ogbonnaya et al. 2001; Ellis et al. 2002). With continuous advances in sequencing technologies, genome-based selection is likely to replace the conventional marker-based genotyping approach to provide a powerful tool for high resolution mapping and large-scale gene discovery. However, genomewide selection requires several important components including (a) very high marker density (with hundreds of thousands of SNP markers) that cannot be met with any of the currently available marker technologies, (b) high throughput low cost DNA extraction method, (c) rapid, cost effective and high throughput large-scale genotyping system, (d) efficient sample collection and tracking system, and (e) automated data management, analyses and interpretation. The future of molecular breeding is therefore building huge data sets and mathematical models that can predict genotypes that will perform well under specific environments. This information will then be used to help breeders create the right cross that most efficiently maximizes agricultural output (yield and quality) with use of minimal input resources.

REFERENCES

AUSTIN, D.F. and LEE, M. Comparative mapping in $F_{2:3}$ and $F_{6:7}$ generations of quantitative trait loci for grain yield and yield component in maize. *TAG Theoretical and Applied Genetics*, May 1996, vol. 92, no. 7, p. 817-826. [CrossRef]

AYOUB, M. and MATHER, D.E. Effectiveness of selective genotyping for detection of quantitative trait loci: an analysis of grain and malt quality traits in three barley populations. *Genome*, August 2002, vol. 45, no. 6, p. 1116-1124. [CrossRef]

BAO, J.S.; WU, Y.R.; HU, B.; WU, P.; CUI, H.R. and SHU, Q.Y. QTL for rice grain quality based on a DH population derived from parents with similar apparent amylose content. *Euphytica*, December 2002, vol. 128, no. 3, p. 317-324. [CrossRef]

BARUA, U.M.; CHALMERS, K.J.; HACKETT, C.A.; THOMAS W.T.; POWELL, W. and WAUGH, R. Identification of RAPD markers linked to a *Rhynchosporium secalis* resistance locus in barley using near-isogenic lines and bulked segregant analysis. *Heredity*, August 1993, vol. 71, no. 2, p. 177-184. [CrossRef]

BASTEN, Christopher J.; WEIR, Bruce S. and ZENG, Zhao-Bang. *QTL Cartographer: A reference manual and tutorial for QTL mapping*. Raleigh, NC, Department of Statistics, North Carolina State University, 2002, [cited on July 2010]. Available from Internet: http://statgen.ncsu.edu/qtlcart/manual/node357.html.

BASTEN, Christopher; WEIR, Bruce S. and ZENG, Zhao-Bang. Zmap -a QTL cartographer. In: SMITH, C.; GAVORA J.S.; BENKEL, B.; CHESNAIS, J.; FAIRFULL, W.; GIBSON, J.P.; KENNEDY, B.W. and BURNSIDE, E.B. eds. *World Congress on Genetics Applied to Livestock Production, 5th, 7th - 12th August, 1994, Guelph, ON., Canada*. Proceedings. Computing strategies and software. Guelph, Organizing Committee, 1994. Vol. 22, p. 65-66.

BEAVIS, W.D. QTL analyses: Power, precision and accuracy. In: PATERSON, A.H. ed. *Molecular Dissection of Complex Traits*. Boca Raton, FL, CRC Press, 1998. p. 145-161.

BEHN, Anita; HARTL, Lorenz; SCHWEIZER, Günther and BAUMER, Max. Molecular mapping of QTLs for non-parasitic leaf spot resistance and comparison of half-sib DH populations in spring barley. *Euphytica*, January 2005, vol. 141, no. 3, p. 291-299. [CrossRef]

BELKHIR, K.; BORSA, P.; CHIKHI, L.; RAUFASTE, N. and BONHOMME, F. *GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations.* Montpellier, Laboratoire Génome, Populations, Interactions, Université de Montpellier II, 1996-2004, [cited on August 2010]. Available from Internet: http://www.genetix.univ-montp2.fr/genetix/genetix.htm.

BERNARDO, Rex. Molecular markers and selection for complex traits in plants: learning from the last 20 years. *Crop Science,* September 2008, vol. 48, no. 5, p. 1649-1664. [CrossRef]

BERNARDO, R. Genomewide selection for rapid introgression of exotic germplasm in maize. *Crop Science*, March 2009, vol. 49, no. 2, p. 419-425. [CrossRef]

BERNARDO, Rex and CHARCOSSET, Alain. Usefulness of gene information in marker-assisted recurrent selection: A simulation appraisal. *Crop Science*, February 2006, vol. 46, no. 2, p. 614-621. [CrossRef]

BERNARDO, Rex and YU, Jianming. Prospects for genomewide selection for quantitative traits in maize. *Crop Scienc*e, May 2007, vol. 47, no. 3, p. 1082-1090. [CrossRef]

BJØRNSTAD, Åsmund; WESTAD, Frank and MARTENS, Harald. Analysis of genetic marker-phenotype relationships by jack-knifed partial least squares regression (PLSR). *Hereditas*, November 2004, vol. 141, no. 2, p. 149-165. [CrossRef]

BOUCHEZ, Agnès; HOSPITAL, Frédéric; CAUSSE, Mathilde; GALLAIS, André and CHARCOSSET, Alain. Marker assisted introgression of favorable alleles at quantitative trait loci between maize elite lines. *Genetics*, December 2002, vol. 162, no. 4, p. 1945-1959.

BRADBURY, Peter J.; ZHANG, Zhiwu; KROON, Dallas E.; CASSTEVENS, Terry M.; RAMDOSS, Yogesh and BUCKLER, Edward S. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, October 2007, vol. 23, no. 19, p. 2633-2635. [CrossRef]

BRADSHAW, H.D. Molecular genetics of populus. In: STETTLER, R.F.; BRADSHAW, H.D.; HEILMAN, P.E. and HINCKLEY, T.M. eds. *Biology of populus and its implications for management and conservation, Part 1*. Ottawa, NRC Research Press, 1996, p. 183-199.

BRAUER, Matthew J.; CHRISTIANSON, Cheryl M.; PAI, Dave A. and DUNHAM, Maitreya J. Mapping novel traits by array-assisted bulk segregant analysis in *Saccharomyces cerevisiae*. *Genetics*, July 2006, vol. 173, no. 3, p. 1813-1816. [CrossRef]

BREEN, G.; SHAM, P.; LI, T.; SHAW, D.; COLLIER, D.A. and SAINT CLAIR, D. Accuracy and sensitivity of DNA pooling with microsatellite repeats using capillary electrophoresis. *Molecular and Cellular Probes*, October 1999, vol. 13, no. 5, p. 359-365. [CrossRef]

BRESEGHELLO, Flavio and SORRELLS, Mark E. Association analysis as a strategy for improvement of quantitative traits in plants. *Crop Science*, June 2006, vol. 46, no. 3, p. 1323-1330. [CrossRef]

BROMAN, K.W.; Review of statistical methods for QTL mapping in experimental crosses. *Laboratory Animals*, 2001, vol. 30, no. 7, p. 44-52.

BROWN, Garth R.; BASSONI, Daniel L.; GILL, Geoffrey P.; FONTANA, Joseph R.; WHEELER, Nicholas C.; MEGRAW, Robert A.; DAVIS, Mark F.; SEWELL, Mitchell M.; TUSKAN, Gerald A. and NEALE, David B. Identification of quantitative trait loci influencing wood property traits in loblolly pine (*Pinus taeda* L.). III. QTL verification and candidate gene mapping. *Genetics*, August 2003, vol. 164, no. 4, p. 1537-1546.

BUCKLER, Edward S.; BROWNE, Chris; ERSOZ, Elhan; FLINT-GARCIA, Sherry; GARCIA, Arturo; HOLLAND, James B.; BRADBURY, Peter J.; ACHARYA, Charlotte B.; BROWN, Patrick J.; GLAUBITZ, Jeffrey C.; GOODMAN, Major M.; HARJES, Carlos; GUILL, Kate; KROON, Dallas E.; LARSSON, Sara; LEPAK, Nicholas K.; LI, Huihui; MITCHELL, Sharon E.; PRESSOIR, Gael; PEIFFER, Jason A.; ROSAS, Marco Oropeza; ROCHEFORD, Torbert R.; ROMAY, M. Cinta; ROMERO, Susan; SALVO, Stella; VILLEDA, Hector Sanchez; DA SILVA, H. Sofia; SUN, Qi; TIAN, Feng; UPADYAYULA, Narasimham; WARW, Doreen; YATES, Heather; YU, Jianming; ZHANG, Zhiwu; KRESOVICH, Stephen and McMULLEN, Michael D. The genetic architecture of maize flowering time. *Science*, August 2009, vol. 325, no. 5941, p. 714-718. [CrossRef]

CALDWELL, Katherine S.; RUSSELL, Joanne; LANGRIDGE, Peter and POWELL, Wayne. Extreme population-dependent linkage disequilibrium detected in an inbreeding plant species, *Hordeum vulgare*. *Genetics*, January 2006, vol. 172, no. 1, p. 557-567. [CrossRef]

CARBONELL, E.A.; ASINS, M.A.; BASELGA, M.; BALANSARD, E. and GERIG, T.M. Power studies in the estimation of genetic parameters of quantitative trait loci for backcross and doubled haploid populations. *TAG Theoretical Applied Genetics*, May 1993, vol. 86, no. 4, p. 411-416. [CrossRef]

CATO, S.; GARDNER, R.; KENT, J. and RICHARDSON, T. A rapid PCR-based method for genetically mapping ESTs. *TAG Theoretical Applied Genetics*, February 2001, vol. 102, no. 2-3, p. 296-306. [CrossRef]

CHAPMAN, A.; PANTALONE, V.R.; USTUN, A.; ALLEN, F.L.; LANDAU-ELLIS, D.; TRIGIANO, R.N. and GRESSHOF, P.M. Quantitative trait loci for agronomic and seed quality traits in an F_2 and $F_{4:6}$ soybean population. *Euphytica*, February 2003, vol. 129, no. 3, p. 387-393. [CrossRef]

CHARCOSSET, A. and GALLAIS, A. Estimation of the contribution of quantitative trait loci (QTL) to the variance of quantitative trait by means of genetic markers. *TAG Theoretical Applied Genetics*, December 1996, vol. 93, no. 8, p. 1193-1201. [CrossRef]

CHASE, K.; ADLER, F.R. and LARK, K.G. Epistat: a computer program for identifying and testing interaction between pairs of quantitative trait loci. *TAG Theoretical and Applied Genetics*, June 1997, vol. 94, no. 6-7, p. 724-730.[CrossRef]

CHASE, Kevin; CARRIER, David R.; ADLER, Frederick R.; JARVIK, Tyler; OSTRANDER, Elaine A.; LORENTZEN, Travis D. and LARK, Karl G. Genetic basis for systems of skeletal quantitative traits: principal component analysis of the canid skeleton. *Proceeding National Academy Sciences of the United States of America*, 2002, vol. 99, no. 15, p. 9930-9935. [CrossRef]

CHING, Ada; CALDWELL, Katherine S.; JUNG, Mark; DOLAN, Maurine; SMITH, Oscar S. (Howie); TINGEY, Scott; MORGANTE, Michele and RAFALSKI, Antoni J. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genetics*, October 2002, vol. 3, no. 19. [CrossRef]

CHURCHILL, G.A. and DOERGE, R.W. Empirical threshold values for quantitative trait mapping. *Genetics*, November 1994, vol.138, no. 3, p. 963-971.

COLLARD, B.; JAHUFER M.Z.Z.; BROUWER, J.B. and PANG, E.C.K. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop

improvement: the basic concepts. *Euphytica*, January 2005, vol. 142, no. 1-2, p. 169-196. [CrossRef]

COLLARD, Bertrand; MACE, Emma; McPHAIL, Mark; WENZL, Peter; CAKIR, Mehmet; FOX, Glen; POULSEN, David and JORDAN, David. How accurate are the marker orders in crop linkage maps generated from large marker datasets? *Crop and Pasture Science*, 2009, vol. 60, no. 4, p. 362-372. [CrossRef]

COMAI, Luca; YOUNG, Kim; TILL, Bradley J.; REYNOLDS, Steven H.; GREENE, Elizabeth A.; CODOMO, Christine A.; ENNS, Linda C.; JOHNSON, Jessica E.; BURTNER, Chris; ODDEN, Anthony R. and HENIKOFF, Steven. Efficient discovery of DNA polymorphisms in natural populations by EcoTILLING. *The Plant Journal*, March 2004, vol. 37, no. 5, p. 778-786. [CrossRef]

CONCIBIDO, Vergel C.; DIERS, Brian W. and ARELLI, Prakash R. A decade of QTL mapping for cyst nematode resistance in soybean. *Crop Science*, July 2004, vol. 44, no. 4, p. 1121-113.

COPPIETERS, Wouter; KVASZ, Alexandre; FARNIR, Frédéric; ARRANZ, Juan-Jose; GRISART, Bernard; MACKINNON, Margaret and GEORGES, Michel. A rank-based non parametric method to map QTL in outbred half-sib pedigrees: application to milk production in a grand-daughter design. *Genetics*, July 1998, vol. 149, no. 3, p. 1547-1555.

DARVASI, Ariel. The effect of selective genotyping on QTL mapping accuracy. *Mammalian Genome*, January 1997, vol. 8, no. 1, p. 67-68. [CrossRef]

DARVASI, Ariel. Experimental strategies for the genetic dissection of complex traits in animal models. *Nature Genetics*, January 1998, vol. 18, no.1, p. 19-24. [CrossRef]

DARVASI, A. and SOLLER, M. Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. *TAG Theoretical and Applied Genetics*, November 1992, vol. 85, no. 2-3, p. 353-359. [CrossRef]

DARVASI, A.; WEINREB, A.; MINKE, V.; WELLER, J.I. and SOLLER, M. Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. *Genetics*, July 1993, vol. 134, no. 3, p. 943-951.

DARVASI, A. and SOLLER, M. Selective DNA pooling for determination of linkage between a molecular marker and a quantitative trait locus. *Genetics,* December 1994, vol.138, no 4, p. 1365-1373.

DARVASI, A. and SOLLER, M. Advanced intercross lines, an experimental population for fine genetic mapping. *Genetics,* November 1995, vol. 141, no. 3, p. 1199-1207.

DEVLIN, B. and ROEDER, K. Genomic control for association studies. *Biometrics*, December 1999, vol. 55, no. 4, p. 997-1004. [CrossRef]

D'HOOP, Björn B.; PAULO, Maria João; MANK, Rolf A.; VAN ECK, Herman J. and VAN EEUWIJK, Fred A. Association mapping of quality traits in potato (*Solanum tuberosum* L.). Euphytica, May 2008, vol. 161, no. 1-2,p. 47-60.

DOERGE, R.W. and CHURCHILL, G.A. Permutation tests for multiple loci affecting a quantitative character. *Genetics*, January 1996, vol. 142, no. 1, p. 285-294.

DOERGE, Rebecca W. Multifactorial genetics: Mapping and analysis of quantitative trait loci in experimental populations. *Nature Review Genetics*, January 2002, vol. 3, no. 1, p. 43-52. [CrossRef]

EAST, E.M. Studies on size inheritance in nicotiana. *Genetics*, March 1916, vol.1, no. 2, p. 164-176.

EDWARDS, M.D.; STUBER, C.W. and WENDEL, J.F. Molecular-marker-facilitated investigations of quantitative-trait loci in maize. I. Numbers, genomic distribution and types of gene action. *Genetics*, May 1987, no. 116:p. 113-125.

EFRON, B. Bootstrap methods: another look at the jackknife. *Annals of Statistics*, January 1979, vol. 7, no. 1, p. 1-26. [CrossRef]

ELLIS, M.H.; SPEILMEYER, W.; GALE K.R.; REBETZKE G.J. and RICHARDS, R.''Perfect'' markers for the *Rht-B1b* and *Rht-D1b* dwarfing genes in wheat. TAG *Theoretical and Applied Genetics*, November 2002, vol. 105, no. 6-7, p. 1038-1042. [CrossRef]

ESHED, Y. and ZAMIR, D. An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. *Genetics*, November 1995, vol. 141, no. 3, p.1147-1162.

EWENS, Warren J. and SPIELMAN, Richard S. Locating genes by linkage and association. *Theoretical Population Biology*, November 2001, vol. 60, no.3, p.135-139. [CrossRef]

EXCOFFIER, Laurent; LAVAL, Guillaume and SCHNEIDER, Stefan. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, January 2005, vol. 1, p. 47-50.

FALCONER, Douglas S. and MACKAY, Trudy F. C. *Introduction to quantitative genetics*. 4th ed. Harlow, UK, Longman Group, 1996. 464 p. ISBN 978-0-582-24302-6

FALUSH, Daniel; STEPHENS, Matthew and PRITCHARD, Jonathan K. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics,* August 2003, vol.164, p. 1567-1587.

FISCH, R.D.; RAGOT, M. and GAY, G. A generalization of the mixture model in the mapping of quantitative trait loci for progeny from a biparental cross of inbred lines. *Genetics,* May 1996, vol.143, no. 1, p. 571-577.

FLINT-GARCIA, Sherry A.; THORNSBERRY, Jeffry M.; EDWARD, S. and BUCKLER IV. Structure of linkage disequilibrium in plants. *Annual Review of Plant Biology,* June 2003, vol. 54, p. 357-374. [CrossRef]

FOOLAD, M.R.; STOLTZ, T.; DERVINIS, C.; RODRIGUEZ, R.L. and JONES, R.A. Mapping QTLs conferring salt tolerance during germination in tomato by selective genotyping. *Molecular Breeding,* August 1997, vol. 3, no. 4, p. 269-277. [CrossRef]

FOOLAD, M.R.; ZHANG, L.P. and LIN, G.Y. Identification and validation of QTLs for salt tolerance during vegetative growth in tomato by selective genotyping. *Genome*, June 2001, vol. 44, no. 3, p. 444-454. [CrossRef]

FU, Yong-Bi. Applications of bulking in molecular characterization of plant germplasm: a critical review. *Plant Genetic Resources,* August 2003, vol. 1, p. 161-167.

GALLAIS, A.; MOREAU, L. and CHARCOSSET, A. Detection of marker-QTL associations by studying change in marker frequencies with selection. *TAG Theoretical and Applied Genetics*, February 2007, vol. 114, no. 4, p. 669-681. [CrossRef]

GARRIS, Amanda J.; McCOUCH, Susan R. and KRESOVICH, Stephen. Population structure and its effect on haplotype diversity and linkage disequilibrium surrounding the *xa5* locus of rice (Oryza sativa L.). *Genetics,* October 2003, vol. 165, no. 2, p. 759-769.

GAUT, Brandon S. and LONG, Anthony D. The lowdown on linkage disequilibrium. *The Plant Cell*, July 2003, vol. 15, no. 7, p. 1502-1506. [CrossRef]

GILBERT, Hélène and LE ROY, Pascale. Comparison of three multitrait methods for QTL detection. *Genetics Selection Evolution,* May 2003, vol.35, p. 281-304. [CrossRef]

GIOVANONI, James J.; WING, Rod A.; GANAL, Martin W. and TANKSLEY, Steven D. Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. *Nucleic Acids Research,* December 1991, vol. 19, no. 23, p. 6553-6558. [CrossRef]

GOUDET, Jérôme. FSTAT version 1.2: A computer program to calculate F-statistics. *Journal of Heredity,* November 1995, vol. 86, no. 6, p. 485-486.

GRATTAPAGLIA, Dario and SEDEROFF, Ronald. Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. *Genetics*, August 1994, vol. 137, no. 4, p. 1121-1137.

GUPTA, P.K.; ROY, J.K. and PRASAD, M. Single nucleotide polymorphisms: a new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. *Current Science,* February 2001, vol. 80, no. 4, p. 524-535.

GUPTA, Pushpendra K.; RUSTGI, Sachin and KULWAL, Pawan L. Linkage disequilibrium and association studies in higher plants: Present status and future prospects. *Plant Molecular Biology,* March 2005, vol. 57, no. 4, p. 461-485. [CrossRef]

HACKETT, Christine A. Statistical methods for QTL mapping in cereals. *Plant Molecular Biology,* March 2002, vol. 48, no. 5-6, p. 585-599. [CrossRef]

HALEY, C.S. and ANDERSSON, L. Linkage mapping of quantitative trait loci in plants and animals. In: DEAR, Paul H., ed. *Genome mapping: a practical approach*. New York, Oxford University Press, 1997, p. 49-71.

HALEY, C.S. and KNOTT, S.A. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity*, October 1992, vol. 69, no. 4, p. 315-324.

HAMBLIN, Martha T., SALAS FERNANDEZ, Maria G., CASA, Alexandra M., MITCHELL, Sharon E., PATERSON, Andrew H. and Kresovich, Stephen. Equilibrium processes cannot explain high levels of short- and medium-range linkage disequilibrium in the domesticated grass *Sorghum bicolor*. *Genetics,* November 2005, vol.171, no. 3, p. 1247-1256. [CrossRef]

HARRIS, Karen; SUBUDHI, P.K.; BORRELL, Andrew; JORDAN, David; ROSENOW, Darrell; NGUYEN, Henry; KLEIN, Patricia; KLEIN, Robert and MULLET, John. Sorghum stay-green QTL individually reduce post-flowering drought-induced leaf senescence. *Journal of Experimental Botany,* January 2007, vol. 58, no. 2, p. 327-338. [CrossRef]

HE, P.; LI, J.Z.; ZHENG, X.W.; SHEN, L.S.; LU, C.F.; CHEN, Y. and ZHU, L.H. Comparison of molecular linkage maps and agronomic trait loci between DH and RIL populations derived from the same rice cross. *Crop Science*, July 2001, vol., 41, no. 4, p. 1240-1246.

HEFFNER, Elliot L.; SORRELLS, Mark E. and JANNINK, Jean-Luc. Genomic selection for crop improvement. *Crop Science*, January-February 2009, vol. 49, no. 1, p. 1-12. [CrossRef]

HOESCHELE, I.; UIMARI, P.; GRIGNOLA, E.F.; ZHANG, Q. and GAGE, K.M. Advances in statistical methods to map quantitative trait loci in outbred populations. *Genetics,* November 1997, vol. 147, no. 3, p. 1445-1457.

HOLLAND, James B. Genetic architecture of complex traits in plants. *Current Opinion in Plant Biology*, 2007, vol. 10, p. 156-161.

HOSPITAL, F.; MOREAU, L.; LACOUDRE, F.; CHARCOSSET, A. and GALLAIS, A. More on the efficiency of marker assisted selection. *TAG Theoretical and Applied Genetics*, December 1997, vol. 95, no. 8, p. 1181-1189. [CrossRef]

IHAKA, Ross and GENTLEMAN, Robert. R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 1996, vol. 5, no. 3, p. 299-314.

JANNINK, Jean-Luc; BINK, Marco C. and JANSEN, Ritsert C. Using complex plant pedigrees to map valuable genes. *Trends Plant Science*, August 2001, vol. 6, no. 8, p. 337-342. [CrossRef]

JANNINK, Jean-Luc; LORENZ, Aaron J. and IWATA, Hiroyoshi. Genomic selection in plant breeding: from theory to practice. *Briefings in Functional Genomics*, March 2010, vol.9, no. 2, p. 166-177.

JANSEN, Ritsert C. Interval mapping of multiple quantitative trait loci. *Genetics,* September 1993, vol. 135, no. 1, p. 205-211.

JANSEN, Ritsert C.; JOHNSON, David L. and VAN ARENDONK, Johan A.M. A mixture model approach to the mapping of quantitative trait loci in complex populations with an application to multiple cattle families. *Genetics,* January 1998, vol. 148, no. 1, p. 391-400.

JANSEN, Ritsert C. and NAP, Jan-Peter. Genetical genomics: the added value from segregation. *Trends in Genetics,* July 2001, vol. 17, no. 7, p. 388-391. [CrossRef]

JIANG, C.J. and ZENG, Z.B. Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics*, July 1995, vol. 140, no. 3, p. 1111-1127.

JIANG, Changjian Jiang and ZENG, Zhao-Bang. Mapping quantitative trait loci with dominant and missing markers in various crosses from two inbred lines. *Genetica*, June 1997, vol. 101, no. 1, p. 47-58. [CrossRef]

JIN, Chunfang; LAN, Hong; ATTIE, Alan D.; CHURCHILL, Gary A.; BULUTUGLO, Dursun and YANDELL, Brian S. Selective phenotyping for increased efficiency in genetic mapping studies. *Genetics*, December 2004, vol. 168, no. 4, p. 2285-2293. [CrossRef]

JOHNSON, L. Marker assisted sweet corn breeding: A model for specialty crops. In. *Proceedings of the 56th Annual Corn and Sorghum Research Conference. (5-7 December, 2001, Chicago, IL)*. Washington, DC., American Seed Trade Association, 2001. p. 25-30.

JOHNSON, Richard. Marker-assisted selection. In: JANICK, Jules. ed. *Plant Breeding Reviews*, 2004, vol. 24, Part 1. p. 293-309. ISBN 9780470650240. [CrossRef]

JORDE, L.B. Linkage disequilibrium and the search for complex disease genes. *Genome Research*, October 2000, vol. 10, no. 10, p. 1435-1444.[CrossRef]

JUNG, Mark; CHING, Ada; BHATTRAMAKKI, Dinakar; DOLAN, Maureen; TINGEY, Scott; MORGANTE, Michelle and RAFALSKI, Antoni. Linkage disequilibrium and sequence diversity in a 500-kbp region around the *adh1* locus in elite maize germplasm. TAG *Theoretical and Applied Genetics*, August 2004, vol. 109, no. 4, p. 681-689.[CrossRef]

KANTETY, Ramesh V.; ROTA, Mauricio La; MATTHEWS, David E. and SORRELLS, Mark E. Data mining for simple-sequence repeats in expressed sequence tags from barley, maize, rice, sorghum, and wheat. *Plant Molecular Biology*, March 2002, vol. 48, no. 5-6, p. 501-510. [CrossRef]

KAO, Chen Hung; ZENG, Zhao-Bang and TEASDALE, Robert D. Multiple interval mapping for quantitative trait loci. *Genetics,* July 1999, vol. 152, no. 3, p. 1203-1216.

KEARSEY, M.J. and FARQUHAR, A.G.L. QTL analysis; where are we now? *Heredity,* February 1998, vol. 80, no. 2, p. 137-142. [CrossRef]

KNAPP, S.J.; BRIDGES, W.C. and BIRKES, D. Mapping quantitative trait loci using molecular marker linkage maps. *TAG Theoretical and Applied Genetics*, May 1990, vol. 79, no. 5, p. 583-592. [CrossRef]

KNOTT, Sara A. and HALEY, Chris S. Maximum likelihood mapping of quantitative trait loci using full-sib families. *Genetics*, December 1992, vol. 132, no. 4, p. 1211-1222.

KNOTT, S.A. and HALEY, C.S. Multiple least squares for quantitative trait loci detection. *Genetics,* October 2000, vol. 156, no. 2, p. 899-911.

KOEBNER, R. MAS in cereals: Green for maize, amber for rice, still red for wheat and barley. In: *Marker assisted selection: a fast track to increase genetic gain in plant and animal breeding? Session I: MAS in plants.* (17-18 October 2003, Turin, Italy). p. 12-17. Available at http://www.fao.org/biotech/docs/Koebner.pdf (cited July 2009).

KOLB, F.L.; BAI, G.H.; MUEHLBAUER, G.J.; ANDERSON, J.A.; SMITH, K.P. and FEDAK, G. Host plant resistance genes for Fusarium head blight: mapping and manipulation with molecular markers. *Crop Science*, May 2001, vol. 41, no. 3, p. 611-619.

KOORNNEEF, Maarten; ALONSO-BLANCO, Carlos and VREUGDENHILL, Dick. Naturally occurring genetic variation in *Arabidopsis thaliana*. *Annual Review of Plant Biology*, June 2004, vol. 55, p. 141-172. [CrossRef]

KOROL, A.B.; RONIN, Y.I. and KIRZHNER, V.M. Interval mapping of quantitative trait loci employing correlated trait complexes. *Genetics,* July 1995, vol. 140, no. 3, p. 1137-1147.

KROYMANN, Juergen and MITCHELL-OLDS, Thomas. Epistasis and balanced polymorphism influencing complex trait variation. *Nature,* May 2005, vol. 435, no. 7038, p. 95-98. [CrossRef]

KRUGLYAK, Leonid and LANDER, Eric S. A nonparametric approach for mapping quantitative trait loci. *Genetics,* March 1995, vol. 139, no. 3, p. 1421-1428.

KRUGLYAK, Leonid. Prospects for whole genome linkage disequilibrium mapping of common disease genes. *Nature Genetics,* June 1999, vol. 22, no. 2, p. 139-144. [CrossRef]

LANDE, Russell. The response to selection on major and minor mutations affecting a metrical trait. *Heredity*, February 1983, vol. 50, no. 1, p. 47-65. [CrossRef]

LANDER, Eric S. and BOTSTEIN, David. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, January 1989, vol. 121, no. 1, p. 185-199.

LANDER Eric and KRUGLYAK, L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature Genetics*, November 1995, vol. 11, no. 3, p. 241-247. [CrossRef]

LAURIE, Cathy C.; CHASALOW, Scott D.; LEDEAUX, John R.; MCCARROLL, Robert; BUSH, David; HAUGE, Brian; LAI, Chaoqiang; CLARK, Darryl; ROCHEFORD, Torbert R. and DUDLEY, John W. The genetic architecture of response to long-term artificial selection for oil concentration in the maize kernel. *Genetics*, December 2004, vol. 168, no. 4, p. 2141-2155.

LAWSON, Darlene M.; LUNDE, China F. and MUTSCHLER, Martha A. Marker-assisted transfer of acylsugar-mediated pest resistance from the wild tomato, *Lycopersicon pennellii*, to the cultivated tomato, *Lycopersicon esculentum*. *Molecular Breeding*, August 1997, vol. 3, no. 4, p. 307-317. [CrossRef]

LEBOWITZ, R.J.; SOLLER, M. and BECKMANN, J.S. Trait-based analyses for the detection of linkage between marker loci and quantitative trait loci in crosses between inbred lines. *TAG Theoretical and Applied Genetics,* February 1987, vol. 73, no. 4, p. 556-562. [CrossRef]

LECOMTE, L.; DUFFÉ, P.; BURET, M.; SERVIN, B.; HOSPITAL, F. and CAUSSE, M. Marker-assisted introgression of five QTLs controlling fruit quality traits into three tomato lines revealed interactions between QTLs and genetic backgrounds. *TAG Theoretical and Applied Genetics,* August 2004, vol. 109, no. 3, p. 658-668. [CrossRef]

LIN, J.-Z. and RITLAND, K. The effects of selective genotyping on estimates of proportion of recombination between linked quantitative trait loci. *TAG Theoretical and Applied Genetics*, December 1996, vol. 93, no. 8, p. 1261-1266. [CrossRef]

LINCOLN, S., DALY, M. and LANDER, E. *Mapping genes controlling quantitative traits with MAPMAKER/QTL. Version 1.1: A tutorial and reference manual.* 2nd ed. Cambridge, MA., Whitehead Institute Technical Report, 1992. 46 p.

LIU, Jiangjung; MERCER, John M.; STAM, Lynn F.; GIBSON, Gregory C.; ZENG, Zhao Bang and LAURIE, Cathy C. Genetic analysis of a morphological shape difference in the male genitalia of *Drosophila simulans* and *D. mauritiana*. *Genetics*, April 1996, vol. 142, no. 4, p. 1129-1145.

LIU, Z. and DEKKERS, J.C.M. Least squares interval mapping of quantitative trait loci under the infinitesimal genetic model in outbred populations. *Genetics,* January 1998, vol. 148, no. 1, p. 495-505.

LU, Quing; CUI, Yuehua and WU, Rongling. A multilocus likelihood approach to joint modeling of linkage, parental diplotype and gene order in a full-sib family. *BMC Genetics*, July 2004, vol. 5. [CrossRef]

LUND, M.S.; SORENSEN, P.; GULDBRANDTSEN, B. and SORENSEN, D.A. Multitrait fine mapping of quantitative trait loci using combined linkage disequilibria and linkage analysis. *Genetics*, January 2003, vol. 163, no. 1, p. 405-410.

LYNCH, M. and WALSH, B. *Genetic and analysis of quantitative traits.* Sunderland, MA., Sinauer Associates, 1998, 980 p. ISBN 0-87893-481-2.

MA, C.X.; LIN, M.; LITTELL, R.C.; YIN, T. and WU, R. A likelihood approach for mapping growth trajectories using dominant markers in a phase-unknown full-sib family. *TAG Theoretical and Applied Genetics*, February 2004, vol. 108, no. 4, p. 699-705. [CrossRef]

MACKAY, Trudy F.C. The genetic architecture of quantitative traits. *Annual Review of Genetics,* December 2001, vol. 35, p. 303-339. [CrossRef]

MACKAY, Ian and POWELL, Wayne. Methods for linkage disequilibrium mapping in crops. *Trends in Plant Science*, February 2007, vol. 12, no. 2, p. 57-63. [CrossRef]

MACKINNON, M.J. and WELLER, J.I. Methodology and accuracy of estimation of quantitative trait loci parameters in a half-sib design using maximum likelihood. *Genetics,* October 1995, vol. 141, no. 2, p. 755-770.

MAHMOOD, T.; EKUERE, U.; YEH, F.; GOOD, A.G. and STRINGAM, G.R. RFLP linkage analysis and mapping genes controlling the fatty acid profile of *Brassica juncea* using reciprocal DH populations. *TAG Theoretical and Applied Genetics*, July 2003, vol. 107, no. 2, p. 283-290. [CrossRef]

MALIEPAARD, C.; ALSTON, F. H.; VAN ARKEL, G.; BROWN, L. M.; CHEVREAU, E.; DUNEMANN, F.; EVANS, K.M.; GARDINER, S.; GUILFORD, P.; VAN HEUSDEN, A. W.; JANSE, J.; LAURENS, F.; LYNN, J.R.; MANGANARIS, A.G.; DEN NIJS, A.P.M.; PERIAM, N.; RIKKERINK, E.;ROCHE, P.;RYDER, C.; SANSAVINI, S.; SCHMIDT, H.; TARTARINI, S.; VERHAEGH, J.J.; VRIELINK-VAN GINKEL, M. and KING, G. J. Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. *TAG Theoretical and Applied Genetics,* July 1998, vol. 97, no. 1-2, p. 60-73. [CrossRef]

MAMMEN, Enno. Bootstrap and wild bootstrap for high dimensional linear models. *Annals of Statistics*, 1993, vol. 21, no. 1, p. 255-285.

MANLY, Kenneth F. and OLSON, Jane M. Overview of QTL mapping software and introduction to Map Manager QT. *Mammalian Genome*, April 1999, vol. 10, no. 4, p. 327-334. [CrossRef]

MANLY, Kenneth F.; CUDMORE, Robert H. and MEER, Jane M. Map manager QTX, cross-platform software for genetic mapping. *Mammalian Genome*, December 2001 vol. 12, no. 12, p. 930-932. [CrossRef]

MARTINEZ, O. and CURNOW, R.N. Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. *TAG Theoretical and Applied Genetics*, December 1992, vol. 85, no. 4, p. 480-488. [CrossRef]

MARTINEZ, Octavio. Spurious linkage between markers in QTL mapping. *Molecular Breeding*, December 1996, vol. 2, no. 4, p. 351-358. [CrossRef]

MATHER, K. and JINKS, J.L. *Biometrical genetics*. 2nd edition. London, Chapman & Hall, 1982. 382 p.

MATHER, Kristie A.; CAICEDO, Ana L.; POLATO, Nicholas R.; OLSEN, Kenneth M.; MCCOUCH, Susan and PURUGGANAN, Michael D. The extent of linkage disequilibrium in rice (*Oryza sativa L.*). *Genetics*, December 2007, vol. 177, no. 4, p. 2223-2232. [CrossRef]

MAURICIO, Rodney. Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology. *Nature Reviews Genetics*, May 2001, vol. 2, no. 5, p. 370-381. [CrossRef]

McCALLUM, Claire M.; COMAI, Luca; GREENE, Elizabeth A. and HENIKOFF, Steven. Targeted screening for induced mutations. *Nature Biotechnology*, April 2000, vol. 18, no. 4, p. 455-457. [CrossRef]

MELCHINGER, Albrecht E.; UTZ, H. Friedrich and SCHÖN, Chris C. Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. *Genetics*, May 1998, vol. 149, no. 1, p. 383-403.

MELCHINGER, A.E.; UTZ, H.F. and SCHÖN, C.C. QTL analyses of complex traits with cross validation, bootstrapping and other biometric methods. *Euphytica*, June 2004, vol. 137, no. 1, p. 1-11. [CrossRef]

MEUWISSEN, T.H.E. and GODDARD, M.E. Fine mapping of quantitative trait loci using linkage disequilibria with closely linked marker loci. *Genetics*, May 2000, vol. 155, no. 1, p. 421-430.

MEUWISSEN, T.H.E.; HAYES, B.J. and GODDARD, M.E. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, April 2001, vol. 157, no. 4, p. 1819-1829.

MEUWISSEN, Theo H.E. and GODDARD, Mike E. Mapping multiple QTL using linkage disequilibrium and linkage analysis information and multitrait data. *Genetics Selection Evolution*, May-June 2004,vol. 36, no. 2, p. 261-279. [CrossRef]

MICHELMORE, R.; PARAN, I. and KESSELI, R. Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Sciences of the Unated States of America,* November 1991, vol. 88, no. 21, p. 9828-9832. [CrossRef]

MICIC, Z.; HAHN, V.; BAUER, E.; MELCHINGER, A.E.; KNAPP, S.J.; TANG, S. and SCHÖN, C.C. Identification and validation of QTL for *Sclerotinia midstalk* rot resistance in sunflower by selective genotyping. *TAG Theoretical and Applied Genetics*, July 2005, vol. 111, no. 2, p. 233-242. [CrossRef]

MONNA, L.; LIN, H.; KOJIMA, S.; SASAKI, T. and YANO, M. Genetic dissection of a genomic region for a quantitative trait locus, Hd3, into two loci, Hd3a and Hd3b, controlling heading date in rice. *TAG Theoretical and Applied Genetics*, April 2002, vol. 104, no. 5, p. 772-778. [CrossRef]

MOOSE, Stephen P. and MUMM, Rita H. Molecular plant breeding as the foundation for 21st Century crop improvement. *Plant Physiology*, June 2008, vol. 147, no. 3, p. 969-977. [CrossRef]

MOREAU, Laurence; CHARCOSSET, Alain; HOSPITAL, Frédéric and GALLAIS, André. Marker-assisted selection efficiency in populations of finite size. *Genetics,* March 1998*,* vol. 148, no. 3, p. 1353-1365.

NAVABI, Alizera; MATHER, D.E.; BERNIER, J., SPANER, D.M. and ATLIN, G.N. QTL detection with bidirectional and unidirectional selective genotyping: marker-based and trait-based analyses. *TAG Theoretical and Applied Genetics*, January 2009*,* vol. 118, no. 2, p. 347-358. [CrossRef]

NEI, Masatoshi and LI, Wen-Hsiung. Linkage disequilibrium in subdivided populations. *Genetics,* September 1973, vol. 75, no. 1, p. 213-219.

NELSON, James C. QGENE: software for marker-based genomic analysis and breeding. *Molecular Breeding*, June 1997, vol. 3, no. 3, p. 239-245. [CrossRef]

NILLSON-EHLE, H. Kreuzunguntersuchungen an Hafer und Weizen. *Acta Universitatis Lundensis*, 1909, vol. 5, no. 2, p. 1-122.

NORDBORG, Magnus. Linkage disequilibrium, gene trees and selfing: an ancestral recombination graph with partial self fertilization. *Genetics,* February 2000, vol. 154, no. 2, p. 923-929.

NORDBORG, Magnus and WEIGEL, Detlef. Next-generation genetics in plants. *Nature,* December 2008, vol. 456, no. 7223, p. 720-723. [CrossRef]

OGBONNAYA, F.C.; SUBRAHMANYAM, N.C.; MOULLET, O.; DE MAJNIK, J.; EAGLES, H.A.; BROWN, J.S.; EASTWOOD, R.F.; KOLLMORGEN, J. APPELS, R. and LAGUDAH, E.S. Diagnostic DNA markers for cereal cyst nematode resistance in bread wheat. *Australian Journal of Agricultural Research,* 2001 vol. 52, no. 12, p. 1367-1374. [CrossRef]

PALAISA, Kelly A.; MORGANTE, Michele; WILLIAMS, Mark and RAFALSKI, Antoni. Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. *Plant Cell*, August 2003, vol. 15, no. 8, p. 1795-1806. [CrossRef]

PALAISA, Kelly; MORGANTE, Michele; TINGEY, Scott and RAFALSKI, Antoni. Long-range patterns of diversity and linkage disequilibrium surrounding the maize *Y1* gene are indicative of an asymmetric selective sweep. *Proceedings of the National Academy of Sciences* of the United States of America, June 2004, vol. 101, no. 26, p. 9885-9890. [CrossRef]

PATERSON, Andrew H. Making genetic maps. In: *Genome mapping in plants (Biotechnology intelligence unit).* San Diego, CA., Academic Press, 1996, p. 23-39. ISBN 978-0125465908

PATERSON, Andrew H.; LANDER, Eric S.; HEWITT, John D.; PETERSON, Susan; LINCOLN, Stephen E. and TANKSLEY, Steven D. Resolution of quantitative traits into Mendelian factors using a complete linkage map of restriction fragment length polymorphisms. *Nature,* October 1988, vol. 335, no. 6192, p. 721-726. [CrossRef]

PATERSON, A.H.; DAMON, S.; HEWITT, J.D.; ZAMIR, D.; RABINOWITCH, H.D.; LINCOLN, S.E.; LANDER, E.S. and TANKSLEY, S.D. Mendelian factors underlying quantitative traits in tomato: comparison across species, generations, and environments. *Genetics*, January 1991, vol. 127, no. 1, p. 181-197.

PEREZ-ENCISO, M. Sequential bulked typing: a rapid approach for detecting QTLs. *TAG Theoretical and Applied Genetics,* March 1998, vol. 96, no. 3-4, p. 551-557. [CrossRef]

PFLIEGER, Stéphanie; LEFEBVRE, Véronique and CAUSSE, Mathilde. The candidate gene approach in plant genetics: a review. *Molecular Breeding,* December 2001, vol. 7, no. 4, p. 275-291. [CrossRef]

PRASAD, M.; VARSHNEY, R.K.; KUMAR, A.; BALYAN, H.S.; SHARMA P.C., EDWARDS, K.J.; SINGH, H.; DHALIWAL, H.S.; ROY, J.K. and GUPTA, P.K. A microsatellite marker associated with a QTL for grain protein content on chromosome arm 2DL of bread wheat. *TAG Theoretical and Applied Genetics,* July 1999, vol. 99, no. 1-2, p. 341-345. [CrossRef]

PRICE, Alkes L.; PATTERSON, Nick J.; PLENGE, Robert M.; WEINBLATT, Michael E.; SHADICK, Nancy A. and REICH, David. Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics,* August 2006, vol. 38, no. 8, p. 904-909. [CrossRef]

PRITCHARD, Jonathan K. and ROSENBERG, Noah A. Use of unlinked genetic markers to detect population stratification in association studies. *The American Journal of Human Genetics*, July 1999, vol. 65, no. 1, p. 220-228. [CrossRef]

PRITCHARD, Jonathan K.; STEPHENS, Matthew and DONNELLY, Peter. Inference of population structure using multilocus genotype data. *Genetics,* June 2000, vol. 155, no. 2, p. 945-959.

PRITCHARD, Jonathan K. and PRZEWORSKI, Molly. Linkage disequilibrium in humans: models and data. *The American Journal of Human Genetics*, July 2001, vol. 69, no. 1, p. 1-14. [CrossRef]

PUMPHREY, Michael O.; BERNARDO, Rex and ANDERSON, James A. Validating the *Fhb1* QTL for fusarium head blight resistance in near-isogenic wheat lines developed from breeding populations. *Crop Science,* February 2007, vol. 47, no. 1, p. 200-206. [CrossRef]

QUARRIE, Steve A.; LAZIC-JANCIC, Vesna; KOVACEVIC, Dragan; STEED, Andy and PEKIC, Sofija. Bulk segregant analysis with molecular markers and its use for improving drought resistance in maize. *Journal of Experimental Botany*, August 1999, vol. 50, no. 337, p. 1299-1306.

RAFALSKI, Antoni. Applications of single nucleotide polymorphisms in crop genetics. *Current Opinion in Plant Biology,* April 2002, vol. 5, no. 2, p. 94-100. [CrossRef]

RAYMOND, M. and ROUSSET, F. GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity*, May 1995, vol. 86, no. 3, p. 248-249.

REMINGTON, David L.; THORNSBERRY, Jeffry M.; MATSUOKA, Yoshihiro; WILSON, Larissa M.; WHITT, Sherry R.; DOEBLEY, John; KRESOVICH, Stephen; GOODMAN, Major M. and BUCKLER IV, Edward S. Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proceedings of the National Academy of Sciences of the United States of America*, September 2001, vol. 98, no. 20, p. 11479-11484. [CrossRef]

RIBAUT, J.-M.; BANZIGER, M.; BETRAN, J.; JIANG, C.; EDMEADES, G.O.; DREHER, K. and HOISINGTON, D. Use of molecular markers in plant breeding: drought tolerance improvement in tropical maize. In: KANG, Manjit S. ed. *Quantitative genetics, genomics, and plant breeding*. Wallingford, CABI, 2002a. p. 85-99. ISBN 9780851996011.

RIBAUT, J.-M.; JIANG, C. and HOISINGTON, D. Simulation experiments on efficiencies of gene introgression by backcrossing. *Crop Science*, March-April 2002b, vol. 42, no. 2, p. 557-565.

RISCH, Neil and MERIKANGAS, Kathleen. The future of genetic studies of complex human diseases. *Science*, September 1996, vol. 273, no. 5281, p. 1516-1517. [CrossRef]

RITTER, E. and SALAMINI, F. The calculation of recombination frequencies in crosses of allogamous plant species with applications to linkage mapping. *Genetical Research,* February 1996, vol. 67, no. 1, p. 55-65. [CrossRef]

ROBERTSON, A. The nature of quantitative genetic variation. In: BRINK, R.A. and STYLES, E.D. eds. *Heritage from Mendel*. Madison, WI., University of Wisconsin Press, 1967. p. 265-280.

RON, Micha; KLIGER, David; FELDMESSER, Erich; SEROUSSI, Eyal; EZRA, Ephraim and WELLER, Joel Ira. Multiple quantitative trait locus analysis of bovine chromosome 6 in the Israeli Holstein population by a daughter design. *Genetics*, October 2001, vol. 159, no. 2, p. 727-735.

RONIN, Y.I.; KIRZHNER, V.M. and KOROL, A.B. Linkage between loci of quantitative traits and marker loci: multi-trait analysis with a single marker. *TAG Theoretical and Applied Genetics,* May 1995, vol. 90, no. 6, p. 776-786. [CrossRef]

ROY, J.K.; PRASAD, M; VARSHNEY, R.K.; BALYAN, H.S.; BLAKE, T.K.; DHALIWAL, H.S.; SINGH, H.; EDWARDS, K.J. and GUPTA, P.K. Identification of microsatellites on chromosome 6B and a STS on 7D of bread wheat showing an association with preharvest sprouting tolerance. *TAG Theoretical and Applied Genetics*, July 1999, vol. 99, no. 1-2, p. 336-340. [CrossRef]

SALVI, Silvio and TUBEROSA, Roberto. To clone or not to clone plant QTLs: present and future challenges. *Trends in Plant Science*, December 2005, vol. 10, no. 6, p. 297-304. [CrossRef]

SATAGOPAN, J.M.; YANDELL, B.S.; NEWTON, M.A. and OSBORN, T.C. A Bayesian approach to detect quantitative trait loci using Markov chain Monte Carlo. *Genetics*, October 1996, vol. 144, no. 2, p. 805-816.

SAX, Karl. The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics*, November 1923, vol. 8, no. 6, p. 552-560.

SCHON, Chris C.; UTZ, H. Friedrich; GROH, Susanne; TRUBERG, Bernd; OPENSHAW, Steve and MELCHINGER, Albrecht E. Quantitative trait loci mapping based on resampling in a vast maize testcross experiment and its relevance to quantitative genetics for complex traits. *Genetics*, May 2004, vol. 167, no. 1, p. 485-498.

SCHORK, Nicholas J. Extended multipoint identity-by-descent analysis of human quantitative traits: efficiency, power, and modeling considerations. *The American Journal of Human Genetics*, December 1993, vol. 53, no. 6, p. 1306-1313.

SEMAGN, Kassa; BJØRNSTAD, Åsmund; SKINNES, Helge; MARØY, Anne Guri; TARKEGNE, Yalew and WILLIAM, Manilal. Distribution of DArT, AFLP and SSR markers in a genetic linkage map of a double haploid hexaploid wheat population. *Genome,* May 2006a, vol. 49, no. 5, p. 545-555. [CrossRef]

SEMAGN, K.; BJØRNSTAD, Å. and NDJIONDJOP, M.N. An overview of molecular marker methods for plants: *African Journal of Biotechnology, December* 2006b, vol. 5, no. 25, p. 2540-2568.

SEMAGN, K.; BJØRNSTAD, Å. and NDJIONDJOP, M.N. Progress and prospects of marker assisted backcrossing as a tool in crop breeding programs. *African Journal of Biotechnology*, December 2006c, vol. 5, no. 25, p. 2588-2603.

SEMAGN, K.; BJØRNSTAD, Å. and NDJIONDJOP, M.N. Principles, requirements and prospects of genetic mapping in plants. *African Journal of Biotechnology*, December 2006d, vol. 5, no. 25, p. 2569-2587.

SEMAGN, Kassa; SKINNES, Helge; BJØRNSTAD, Åsmund; MARØY, Anne Guri and TARKEGNE, Yalew. Quantitative trait loci controlling fusarium high blight resistance and low deoxynivalenol content in hexaploid wheat population derived from “Arina” and NK93604. *Crop Science*, January 2007, vol. 47, no.1, p. 294-303. [CrossRef]

SHAM, Pak; BADER, Joel S.; CRAIG, Ian; O’DONOVAN, Michael and OWEN, Michel. DNA pooling: a tool for large-scale association studies. *Nature Review of Genetics*, November 2002, vol. 3, no. 11, p. 862-871. [CrossRef]

SHAO, J. Bootstrap model selection. *Journal of American Statistical Association*, June 1996, vol. 91, no. 434, p. 655-665.

SHARBEL, Timothy F.; HAUBOLD, Bernhard and MITCHELL-OLDS, Thomas. Genetic isolation by distance in *Arabidopsis thaliana*: biogeography and postglacial colonization of Europe. *Molecular Ecology,* December 2000, vol. 9, no. 12, p. 2109-2118. [CrossRef]

SHEN, L.; COURTOIS, B.; MCNALLY, K.L.; ROBIN, S. and LI, Z. Evaluation of near-isogenic lines of rice introgressed with QTLs for root depth through marker-aided selection. *TAG Theoretical and Applied Genetics,* July 2001, vol. 103, no. 1, p. 75-83. [CrossRef]

SHEN, X.; ZHOU, M.; LU, W. and OHM, H. Detection of fusarium head blight resistance QTL in a wheat population using bulked segregant analysis. *TAG Theoretical and Applied Genetics*, April 2003, vol. 106, no. 6, p. 1041-1047. [CrossRef]

SILLANPÄÄ, Mikko J. and ARJAS, Elja. Bayesian mapping of multiple quantitative trait loci from incomplete inbred line cross data. *Genetics*, March 1998, vol. 148, no. 3, p. 1373-1388.

SILLANPÄÄ, Mikko J. and ARJAS, Elja. Bayesian mapping of multiple quantitative trait loci from incomplete outbred offspring data. *Genetics*, April 1999, vol. 151, no. 4, p. 1605-1619.

SKINNES, H.; SEMAGN, K.; TARKEGNE, Y.; MARØY, A.G. and BJØRNSTAD, Å. The inheritance of anther extrusion in hexaploid wheat and its relationship to *Fusarium* head blight resistance and deoxynivalenol content. *Plant Breeding*, April 2010, vol.129, no. 2, p. 149-155. [CrossRef]

SOLLER, M.T.; BRODY, T. and GENIZI, A. On the power of experimental designs for detection of linkage between marker loci and quantitative loci in crosses between inbred lines. *TAG Theoretical and Applied Genetics*, January 1976, vol. 47, no. 1, p. 35-39. [CrossRef]

SOLLER, M. and BECKMANN, J.S. Marker-based mapping of quantitative trait loci using replicated progenies. *TAG Theoretical and Applied Genetics*, August 1990, vol. 80, no. 2, p. 205-208. [CrossRef]

SOMERS, Daryl J.; FEDAK, George and SAVARD, Marc. Molecular mapping of novel genes controlling Fusarium head blight resistance and deoxynivalenol accumulation in spring wheat. *Genome,* August 2003, vol.46, no. 4, p. 555-564. [CrossRef]

STUBER, C.W.; LINCOLN, S.E.; WOLFF, D.W.; HELENTJARIS, T. and LANDER, E.S. Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. *Genetics*, November 1992, vol. 132, no. 3, p. 823-839.

SUN, Yanping; WANG, Jiankang; CROUCH, Jonathan H. and XU, Yunbi. Efficiency of selective genotyping for genetic analysis of complex tarits and potential applications in crop improvement. *Molecular Breeding*, January 28. In press, 2010. [CrossRef]

TANKSLEY, S.D. and NELSON, J.C. Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *TAG Theoretical and Applied Genetics*, February 1996, vol. 92, no. 2, p. 191-203. [CrossRef]

TENAILLON, Maud I.; SAWKINS, Mark C.; LONG, Anthony D.; GAUT, Rebecca L. and DOEBLEY, John F. Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays ssp. mays* L.). *Proceedings of the National Academy of Sciences of the United States of America*, July 2001, vol. 98, no. 16, p. 9161-9166. [CrossRef]

THABUIS, A.; PALLOIX, A.; SERVIN; B.; DAUBÈZE, A.M.; SIGNORET, P.; HOSPITAL, F. and LEFEBVRE; V. Marker assisted introgression of 4 *Phytophthora capsici* resistance QTL alleles into a bell pepper line: validation of additive and epistatic effects. *Molecular Breeding*, August 2004, vol. 14, no. 1, p. 9-20. [CrossRef]

THODAY, J.M. Location of polygenes. *Nature*, July 1961, vol. 191, no. 4786, p. 368-370. [CrossRef]

TUBEROSA, Roberto; SALVI, Silvio; SANGUINETI, Maria Corinna; LANDI, Pierangelo; MACCAFERRI, Marco and CONTI, Sergio. Mapping QTLs regulating morpho-physiological traits and yield: case studies, shortcomings and perspectives in drought-stressed maize. *Annals of Botany*, June 2002, vol. 89, no. 7, p. 941-963. [CrossRef]

TUBEROSA, Roberto and SALVI, Silvio. Genomics-based approaches to improve drought tolerance of crops. *Trends in Plant Science*, August 2006, vol. 11, no. 8, p. 405-412. [CrossRef]

UIMARI, P. and HOESCHELE, I. Mapping-linked quantitative trait loci using Bayesian analysis and Markov chain Monte Carlo algorithms. *Genetics*, June 1997, vol. 146, no. 2, p. 735-743.

UPADYAYULA, N.; WASSOM, J.; BOHN, M.O. and ROCHEFORD, T.R. Quantitative trait loci analysis of phenotypic traits and principal components of maize tassel inflorescence architecture. *TAG Theoretical and Applied Genetics*, November 2006, vol. 113, no. 8, p. 1395-1407. [CrossRef]

UTZ, H. Friedrich; MELCHINGER, Albrecht E. and SCHON, Chris C. Bias and sampling error of the estimated proportion of genotypic variance explained by quantitative trait loci determined from experimental data in maize using cross validation and validation with independent samples. *Genetics*, April 2000, vol. 154, no. 4, p. 1839-1849.

UTZ, H.F. and MELCHINGER, A.E. PLABQTL, a computer program to map QTL, Version 1.2. 2006-06-01. Stuttgart, Institute of Plant Breeding, Seed Science, and Population Genetics, University of Hohenheim, 2003. 33 p.

VAN OOIJEN, J.W. MapQTL 6: Software for the mapping of quantitative trait loci in experimental populations of diploid species. Wageningen, The Netherlands, 2009 [cited on August 2010]. Available from Internet: http://www.kyazma.nl/docs/MQ6Manual.pdf.

VARSHNEY, Rajeev K.; HOISINGTON, David A. and TYAGI, Akhilesh K. Advances in cereal genomics and applications in crop breeding. *Trends in Biotechnology*, November 2006, vol. 24, no. 1, p. 490-499. [CrossRef]

VILLAR, M.; LEFEVRE, F.; BRADSHAW, H.D. and DU-CROS, E.T. Molecular genetics of rust resistance in poplars (*Melampsora larici-populina* Kleb/Populus sp.) by bulked segregant analysis in a 2 x 2 factorial mating design. *Genetics*, May 1996, vol. 143, no. 1, p. 531-536.

VISION, Todd J.; BROWN, Daniel G.; SHMOYS, David B.; DURRETT, Richard T. and TANKSLEY, Steven D. Selective mapping: a strategy for optimizing the construction of high-density linkage maps. *Genetics,* May 2000, vol. 155, no. 1, p. 407-420.

VISSCHER, P.; THOMPSON, R. and HALEY, C. Confidence intervals in QTL mapping by bootstrapping. *Genetics,* June 1996, vol. 143, no. 2, p. 1013-1020.

WALL, Jeffrey D.; ANDOLFATTO, Peter and PRZEWORSKI, Molly. Testing models of selection and demography in *Drosophila simulans*. *Genetics,* September 2002, vol. 162, no. 1, p. 203-216.

WANG, G.L. and PATERSON, A.H. Assessment of DNA pooling strategies for mapping of QTLs. *TAG Theoretical and Applied Genetics*, June 1994, vol. 88, no. 3-4, p. 355-361. [CrossRef]

WANG, S.; BASTEN, C.J. and ZENG, Z.B. Windows QTL Cartographer 2.5. Raleigh, NC., Department of Statistics, North Carolina State University, 2007, [cited on August 2010]. Available from Internet: http://statgen.ncsu.edu/qtlcart/WQTLCart.htm.

WEIR, Bruce S. *Genetic data analysis, Vol. II: Methods for discrete population genetic data.*Sunderland, MA, Sinauer, 1996. 376 p. ISBN 0878939024

WELLER, J.I.; KASHI, Y. and SOLLER, M. Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. *Journal of Dairy Science*, September 1990, vol.73, no. 9, p. 2525-2537. [CrossRef]

WELLER, J.I.; WIGGANS, G.R.; VANRADEN, P.M. and RON, M. Application of a canonical transformation to detection of quantitative trait loci with the aid of genetic markers in a multi-trait experiment. *TAG Theoretical and Applied Genetics,* June 1996, vol. 92, no. 8, p. 998-1002. [CrossRef]

WONG, C.K. and BERNARDO, R. Genomewide selection in oil palm: increasing selection gain per unit time and cost with small populations. *TAG Theoretical and Applied Genetics*, April 2008, vol.116, no. 6, p. 815-824. [CrossRef]

WRIGHT, Sewall. *Genetic and biometric foundations.* In: *Evolution and the genetics of populations. Volume 1.* Chicago, University of Chicago Press, 1968. 471 p.

WU, Rongling and ZENG, Zhao-Bang. Joint linkage and linkage disequilibrium mapping in natural populations. *Genetics,* February 2001, vol. 157, no. 2, p. 899-909.

WU, Rongling; MA, Chang-Xing; PAINTER, Ian and ZENG, Zhao-Bang. Simultaneous maximum likelihood estimation of linkage and linkage phases in outcrossing populations. *Theoretical Population Biology*, May 2002, vol. 61, no. 3, p. 349-363. [CrossRef]

XU, Shizhong. A comment on the simple regression method for interval mapping. *Genetics*, December 1995, vol. 141, no. 4, p. 1657-1659.

XU, Shizhong. Mapping quantitative trait loci using multiple families of line crosses. *Genetics*, January 1998, vol. 148, no. 1, p. 517-524.

XU, Y. Quantitative trait loci: separating, pyramiding, and cloning. *Plant Breeding Review,* 1997, vol. 15, p. 85-139.

XU, Yunbi and CROUCH, Jonathan H. Marker-assisted selection in plant breeding: from publications to practice. *Crop Science,* March 2008, vol. 48, no. 2, p. 391-407. [CrossRef]

XU, Yunbi. *Molecular plant breeding*. Wallingford, UK, CABI, 2010. 736 p.

YAMAMOTO, Kimiko and SASAKI, Takuji. Large-scale EST sequencing in rice. *Plant Molecular Biology*, September 1997, vol. 35, no. 1-2, p. 135-144. [CrossRef]

YAN, Jianbing; SHAH, Trushar; WARBURTON, Marilyn L.; BUCKLER, Edward S.; MCMULLEN, Michael D. and CROUCH, Jonathan. Genetic characterization and linkage disequilibrium

estimation of a global maize collection using SNP markers. *PLoS ONE,* December 2009, vol. 4, no. 12. [CrossRef]

YAN, Weikai; KANG, Manjit S. and KANG, Manjit. *GGE Biplot analysis: A graphical tool for breeders, geneticists and agronomists.* CRC Press, 2003. 271 p. ISBN 0849313384.

YOUNG, N.D. Constructing a plant genetic linkage map with DNA markers, In: RONALD, I.K.V. and PHILLIPS, L., eds. *DNA-based markers in plants.* Dordrecht, Kluwer, 1994, p. 39-57.

YOUNG, N.D. QTL mapping and quantitative disease resistance in plants. *Annual Review of Phytopathology,* September 1996, vol. 34, p. 479-501. [CrossRef]

YOUNG, Nevin Dale. A cautiously optimistic vision for marker-assisted breeding. *Molecular Breeding*, December 1999, vol. 5, no. 6, p. 505-510. [CrossRef]

YU, Jianming; PRESSOIR, Gael; BRIGGS, William H.; BI, Irie Vroh; YAMASAKI, Masanori; DOEBLEY, John F.; MCMULLEN, Michael D.; GAUT, Brandon S.; NIELSEN, Dahlia M.; HOLLAND, James B.; KRESOVICH, Stephen and BUCKLER, Edward S. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics*, February 2006, vol.38, no. 2, p. 203-208. [CrossRef]

YU, Jianming and BUCKLER, Edward S. Genetic association mapping and genome organization of maize. *Current Opinion in Biotechnology*, April 2006, vol. 17, no. 2, p. 155-160. [CrossRef]

ZENG, Zhao-Bang. Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proceedings of the National Academy of Sciences*, December 1993, vol. 90, no. 23, p. 10972-10976. [CrossRef]

ZENG, Zhao-Bang. Precision mapping of quantitative trait loci. *Genetics,* April 1994, vol. 136, no. 4, p. 1457-1468.

ZENG, Zhao-Bang; KAO, Chen-Hung and BASTEN, Christopher J. Estimating the genetic architecture of quantitative traits. *Genetical Research*, December 1999, vol. 74, no. 3, p. 279-289. [CrossRef]

ZENG, Zhao-Bang; LIU, Jiangjun; STAM, Lynn F.; KAO, Chen-Hung; MERCER, John Mercer. Genetic architecture of a morphological shape difference between two Drosophila species. *Genetics*, January 2000, vol. 154, no. 1, p. 299-310.

ZHANG, L.P.; LIN, G.Y.; NIÑO-LIU, D. and FOOLAD, M.R. Mapping QTL conferring early blight (*Alternaria solani*) resistance in a *Lycopersicum esculentum* x *L. hirsutum* cross by selective genotyping. *Molecular Breeding*, August 2003, vol. 12, no. 1, p. 3-19. [CrossRef]

ZHANG, Zhiwu; BRADBURY, Peter J.; KROON, Dallas E.; CASSTEVENS, Terry M. and BUCKLER, Edward S. TASSEL 2.0: a software package for association and diversity analyses in plants and animals. In: Plant & Animal Genomes XIV Conference (January 14-18, 2006, San Diego, CA.). *PAG Archives,* [cited on August 2010]. Available from Internet: http://www.maizegenetics.net/bioinformatics.

ZHONG, Shengqiang; DEKKERS, Jack C.M.; FERNANDO, Rohan L. and JANNINK, Jean-Luc. Factors affecting accuracy from genomic selection in populations derived from multiple inbred lines: a barley case study. *Genetics*, May 2009, vol. 182, no. 1, p. 355-364.

ZHU, Chengsong; GORE, Michael; BUCKLER, Edward S. and YU, Jiangming. Status and prospects of association mapping in plants. *Plant Genome,* July 2008, vol. 1, no. 1, p.5-20. [CrossRef]

How to cite this article:

SEMAGN, K.; BJØRNSTAD Å. and XU Y. The genetic dissection of quantitative traits in crops. *Electronic Journal of Biotechnology*, September 2010, vol. 13, no. 5. http://dx.doi.org/10.2225/vol13-issue5-fulltext-14