ISSN 2277-9663
AGRES - An International e-Journal

Volume 1 Issue 2 April-June,2012

MOLECULAR MARKERS IN PLANT GENOME ANALYSIS: A REVIEW

SHEIKH WASEEM*; ACHARYA, S. AND PATEL, J. B.

S. D. AGRICULTURAL UNIVERSITY, SARDARKRUSHINAGAR, GUJARAT, INDIA

Email: waseems84@gmail.com

The concept of genetic markers is not a new one; Mendel used phenotype based genetic markers in his experiment in the nineteenth century. Later, phenotype based genetic markers for Drosophila led to the establishment of the theory of genetic linkage. A molecular marker is defined as a particular segment of DNA that is representative of the differences at the genome level. Molecular markers may or may not correlate with phenotypic expression of a trait. Molecular markers offer numerous advantages over conventional phenotype based markers as they are stable and detectable in all tissues regardless of growth, differentiation, development, or defense status of the cell are not confounded by the environment, pleiotropic and epistatic effects.

The differences that distinguish one plant from another are encoded in the plant's genetic material, the deoxyribonucleic acid (DNA). DNA is packaged in chromosome pairs, one coming from each parent. The genes, which control a plant's characteristics, are located on specific segments of each chromosome. All of the genes carried by a single gamete is known as genome (King and Stansfield, 1990). Although the whole genome sequence is now available for a few plant species such as *Arabidopsis thaliana* (The Arabidopsis Genome Initiative, 2000) and rice (The Rice Genome Mapping Project, 2005), to help identify specific genes located on a particular chromosome, most scientists use an indirect method called genetic markers. Since the markers and the genes they are close together on the same chromosome, they tend to stay together as each generation of plants is produced. Where markers occur on a chromosome, and how close they are to specific genes, scientists can create a genetic linkage map. Such genetic maps serve several purposes, including detailed analysis of associations between economically important traits and genes or quantitative trait loci (QTLs) and facilitate the

introgression of desirable genes or QTLs through marker-assisted selection (Semagn *et al.,* 2006b).

Genetic markers fall into three broad classes: visually assessable traits (morphological and agronomic traits), gene product (biochemical markers), and DNA assay (molecular markers). The idea of using genetic markers appeared very early in literatures (Sax, 1932; Wexelsen, 1933) but the development of electrophoretic assays of isozymes (Markert and Moller, 1959) and molecular markers (Welsh and McClelland, 1990; Williams *et al.,*1990; Adams *et al.,*1991; Caetano-Anolles *et al.,*1991; Akkaya *et al.,*1992; Akopyanz *et al.,*1992; Jordan and Humphries, 1994; Zietkiewicz *et al.,*1994; Vos *et al.,*1995; Jaccoud *et al*.,2001) have greatly improved the understanding of biological sciences. Molecular markers should not be considered as normal genes, as they usually do not have any biological effect, and instead can be thought of as constant landmarks in the genome.

They are identifiable DNA sequences, found at specific locations of the genome, and transmitted by the standard laws of inheritance from one generation to the next. The existence of various molecular techniques and differences in their principles and methodologies require careful consideration in choosing one or more of such marker types. This review article deals on the basic principles, requirements, and advantages and disadvantages of the most widely used molecular markers for genetic diversity studies, genetic mapping, marker-trait association studies, and marker assisted selection programs.

Molecular Markers

The various molecular markers can be classified into different groups based on:

1. Mode of transmission (biparental nuclear inheritance, maternal nuclear inheritance, maternal organelle inheritance, or paternal organelle inheritance).

2. Mode of gene action (dominant or co-dominant markers).

3. Method of analysis (hybridization-based or PCR based markers).

ISSN 2277-9663
AGRES - An International e-Journal

Volume 1 Issue 2 April-June,2012

RELATIONSHIPS AMONG DNA MARKERS DETERMINES BY THE PRINCIPALS OF GENETICS

Linkage

Genetic mapping is made possible by the fact that the nuclear genome of higher organisms is organized and transmitted as linear units, called chromosomes. Genetic linkage or co-transmission from parent to progeny of genetic markers which are close together on the same chromosome provides a means for determining the order of DNA markers along the chromosome.

Genetic analysis of simple and complex traits

Inheritance of simple Mendelian factors may be complicated by segregation distortion, which is common in sexual progeny of many pathogenic fungi. Segregation distortion causes segregation ratios to deviate from the expected and may prompt investigators to propose unnecessarily complex models for the genetic control of simple traits such as specific virulence. Segregation analysis of genetic markers distributed throughout the genome can pinpoint areas undergoing segregation distortion in specific crosses. Identification of markers linked to the genes controlling the trait of interest will then reveal whether the trait is controlled by a single factor or is under more complex control.

Many, if not most, traits are not controlled by single genes are, effected by the environment to some extent, or are, incompletely penetrant. Genetic analysis of such traits by analysis of co-segregation of molecular markers distributed throughout the genome has many advantages over more classical approaches to quantitative genetics. Classical quantitative genetic analysis can sometimes determine the number of genes governing a trait and the average degree of dominance (in diploids or dikaryons) at the loci, but generally assumes that the genes controlling a trait are roughly equal and additive in their effects. Quantitative trait analysis with molecular markers can identify the magnitude and map position of individual loci. The technique is particularly powerful after a more or less complete genetic map has been constructed and segregation of the whole genome can be assayed as intervals between the various markers. An example of an important and interesting complex trait is host adaptation. The genetic control of host adaptation can

ISSN 2277-9663
AGRES - An International e-Journal

Volume 1 Issue 2 April-June,2012

be an important agricultural consideration. For example, corn and sorghum are grown in close proximity in many areas of the world and share many pathogens, e.g. *C. graminicola, F. moniliforme*, and *Peronosclerospora sorghi*. Experiments with *C. graminicola*, however, have indicated that the isolates adapted to the two species are reproductively isolated. Similarly, host preference is exhibited by members of the different mating populations of *F. moniliforme*.

Map-based cloning

Map based cloning where genes are isolated on the basis of their proximity to DNA markers, is an increasingly popular method of cloning genes with unknown products. Distances must be spanned by chromosome walking once a marker is identified near the gene of interest. As a result, some very interesting genes have been isolated by map based cloning to date.

Physical distances can be estimated from the recombination distance between a DNA marker and a gene, the relationship between the two types of distances is far from absolute. At a more localized level, physical distance may be underestimated near centromeres or rDNA arrays or may be overestimated near recombination hot-spots.

Technological developments continually increase the discriminative power and cost effectiveness of profiling. Detailed and discriminative genetic profiles provide extremely powerful and effective procedures to allow meaningful and valid comparisons among inbred lines, varieties and hybrids to be made with respect to germplasm identification and ownership. These profiles when used in conjunction with pedigree and performance data provide a complete source of information to protect intellectual property rights that relate to plant varieties. DNA sequencing provides ultimate fine scale measurement of differences, since all markers are derived (directly or indirectly) from sequence polymorphisms. It is possible to identify varieties and specific quality attributes in mixtures of grain provided appropriate sequence based assays are available. It is not completely necessary to construct a detailed genetic map for map based cloning if one is interested in only one or a few loci. Polymorphisms using RAPD analysis among the lines should map to the locus which is polymorphic in the isogenic lines. As near isogenic lines took considerable tome to be constructed, bulked segregant analysis has also been used

ISSN 2277-9663
AGRES - An International e-Journal

Volume 1 Issue 2 April-June,2012

with good success. Although map based cloning is the only strategy wholly dependent upon it, genetic mapping is a key component of many gene isolation strategies. Thereby molecular marker data help promote investment, productivity and use of important genes in agriculture.

Marker Assisted Selection in crop plants

Development of high yielding varieties, hybrids and populations is one of the most cherished objectives in breeding of self- as well as cross-pollination crops. This is achieved through concentration of favourable gene assembles in desired agronomic background. In self-pollinated crops the desirable genes are scattered over a multitude of pure lines. In conventional breeding the elite types are crossed and recombinants are reestablished into pure lines with desirable gene constellation concentrated. Likewise in cross-pollination crops, the populations are developed which have heterozygote per se with higher frequency of favourable alleles. The propensity and effectiveness with which amassing favorable alleles is done depends upon how closely phenotypes represent genotype and the selection skills of the plant breeders. As large scale genotype x environmental interaction and environmental effects prevail in actual field condition, selection of elite type is therefore is subjective to some extent. Moreover, uncontrollable environmental factor and g x e interaction caused bias in selection to unknown degree and hence slower the pace of progress in breeding endeavors. Molecular markers being independent of the environment offer potent tool to directly select for the genotype. The efficacy of molecular marker – assisted selection for target traits depends upon the linkage relationship of marker with target trait and resolving power of co-segregating molecular marker i.e. of course the MAS technology should be economically affordable and easy to practice also. This necessitates of intensification of research and development efforts to further refine this technology and training to develop human resource to use MAS in crop improvement. Fortunately, there is perceptive change in attitudes as the plant breeders and biotechnologists now agree to rub shoulders to put MAS to practice. Hopefully coming decades will witness that MAS being used increasingly for the crop improvement.

APPLICATIONS OF MOLECULAR MARKERS IN PLANT BREEDING

The most basic applications of the molecular marker techniques in marker assisted breeding include genetic diversity analysis, variety

ISSN 2277-9663
AGRES - An International e-Journal

Volume 1 Issue 2 April-June,2012

identification, isolation of markers tightly linked to specific genes, and marker assisted back-crossing.

Genotyping, Genetic Diversity and Seed Purity Analysis

Genotyping using DNA markers can be considered as the most reliable method for the identification of lines and varieties. Therefore the DNA fingerprinting methods can be used to analyze the purity of seed lots. Genetic distance analysis can be a powerful tool for breeders to identify different heterotic groups and to increase the efficiency of finding crosses with good specific combinability (SCA). To determine the genetic distance between lines and groups of lines, the lines are fingerprinted and the marker-presence or absence is scored for each line. Based on the obtained score table, similarity indices can be calculated for all combinations of lines. Subsequently, the relatedness amongst the lines can be visualized using a dendrogram display or PCA plots.

Indirect selection

Indirect selection can be an advantageous method of selection in plant breeding. Especially for traits for which the phenotypic tests are unreliable, expensive or destructive to the plant (e.g. root parameters), markers can offer a solution. Before indirect selection can be applied, the genetic basis of the trait of interest needs to be elucidated and markers linked to the gene(s) of interest have to be identified. Once linked markers have been identified, the markers (if non-PCR like RFLP) can be converted into simple PCR assays (like STS), which allow screening of large numbers of plants for the trait of interest in a cost effective manner. A suitable linked DNA marker should allow the prediction of the phenotype in a range of the germplasm. The occurrence of multiple alleles in the germplasm for a desired locus may sometimes complicate the identification of markers with a good predictive value in the germplasm. There are different approaches for the monogenic and polygenic traits.

Monogenic traits

For the identification of markers linked with monogenic traits, different approaches can be followed. The preferable approaches are all

ISSN 2277-9663
AGRES - An International e-Journal

Volume 1 Issue 2 April-June,2012

based on screening a limited number of samples with a relatively large number of primer pairs. This way, many loci can be screened with a limited effort. The number of lanes per fingerprint can be limited by screening on set(s) of Near Isogenic Lines (NIL's), if these are available. Candidate markers that are identified in this way are then screened on a panel of phenotypically well characterized lines (for example male sterile and fertile lines, resistant and susceptible lines etc.) to confirm their linkage and to determine the predictive value of the markers. Another efficient approach consists of the 'Bulked Segregant Analysis' (B.S.A.) method (Michelmore *et al.,* 1991). For this type of screening, individuals from a segregating population are pooled on the basis of their phenotype (contrasting characters), and the pools are then screened until a sufficient number of markers emerges. This method can be used for both dominant and recessive monogenic traits. For dominant genes, 'cis' markers (linked with the trait of interest) will emerge from the screening, whereas 'trans' markers (linked with the opposite allele) will be identified for recessive traits.

Polygenic (quantitative) traits

The classical approach for the identification of loci involved in complex polygenic traits consists in the screening of a large number of individuals from a segregating population with a set of markers that are evenly distributed throughout the genome. Subsequently, statistical analysis is performed to identify regions in the genome that are involved in the trait. The laborious nature of this approach makes it unrealistic to screen sufficiently large populations to precisely locate the quantitative trait loci (QTLs). As a consequence, the QTLs cannot be localized precisely on the map and closely linked markers cannot be obtained, thereby preventing the broad scale application of indirect selection for quantitative traits.

For this reason, a new approach for the identification of QTLs markers, based on the B.S.A. principle, was investigated. With this goal, an oilseed rape F_2 population of » 2500 individuals segregating for two quantitative traits, glucosinolate and erucic acid contents, has been used. Based on the phenotypic scores, bulks were composed and approximately 2000 loci were screened on those bulks using the AFLP fingerprinting technique. Candidate markers that were identified on the bulk screening

ISSN 2277-9663
AGRES - An International e-Journal

Volume 1 Issue 2 April-June,2012

were then analyzed on » 200 randomly chosen individuals of the F_2 population. This screening demonstrated that the candidate markers identified using the B.S.A. screening approach were derived from three different loci involved in glucosinolate content and two QTLs involved in erucic acid content. The results have been confirmed by independent studies in which the same map positions have been identified to be involved in the respective traits (Toroser *et al.,*1995; Jourdren *et al.,*1996). The results demonstrate that a B.S.A. strategy may be useful even for the identification of markers for quantitative traits.

Bulk Segregant Analysis (BSA)

Often a geneticist is not interested in developing a molecular map, but would rather find a few markers that are closely linked to a specific trait. The identification of these markers is often achieved by a procedure called bulk segregant analysis. The essence of this procedure is the creation of a bulk sample of DNA for analysis by pooling DNA from individuals with similar phenotypes. For example, you may be interested in finding a molecular locus linked to a disease resistance locus. You would create two bulk DNA samples, one containing DNA from plants or lines that are resistant to the disease and a second bulk containing DNA from plants or lines that are susceptible to the disease. Each of these bulk DNA samples will contain a random sample of all the loci in the genome, except for those that are in the region of the gene upon which the bulking occurred. Therefore, any difference in RFLP or RAPD pattern between these two bulks should be linked to the locus upon which the bulk was developed. This is a powerful technique that has gained wide acceptance in the few years since it was first described.

Marker Assisted Backcross Breeding

With the cost reductions that can be achieved using marker technology, Marker Assisted Backcross breeding is now at the verge of becoming a standard application in modern plant breeding. Two different aspects can be distinguished in backcross breeding: (i) Selection for high recurrent parent: In this application, the DNA fingerprints are used to calculate the % recurrent parent genome in each backcross individual, hereby taking the genome representation of the markers into account. (ii)

ISSN 2277-9663
AGRES - An International e-Journal

Volume 1 Issue 2 April-June,2012

Selection against linkage drag: when negative characteristics are linked with the trait that needs to be introgressed, molecular markers can be used to select for recombinants in the region. After phenotypic testing of these recombinants, individuals may be selected in which the region responsible for the linkage drag has been removed from the locus of interest.

Gene Pyramiding

Gene pyramiding is a very useful approach for the introgression of genes controlling different agronomic traits to ensure that a variety may simultaneously acquire several traits. For example, genes leading to resistance different races or biotypes to a disease or insect pest can be pyramided together to make a line with multi-race multi-biotype resistances, which could be more durable than any single-race or single-biotype resistance (Jiang *et al.,*2004). The joint expression of pyramided genes was found to provide numerical increases or a broader spectrum of resistance over that conferred by single genes through gene interaction and quantitative complementation (Yoshimura *et al.,*1995; Singh *et al.,*2001). Gene pyramiding has been successfully applied in several crop breeding programs, and many varieties and lines possessing multiple attributes have been produced (Huang *et al.,*1997; Samis *et al.,*2002). Gene pyramiding is, however, difficult using conventional breeding methods due to the dominance and epistasis effects of genes governing disease resistance (the stronger resistance genes will always mask the less strong, which cannot be revealed without screening using a virulent strain on the former – itself undesirable). Moreover, genes with similar reactions to two or more races – so called race-non specific partial resistance are difficult to identify and transfer through conventional approaches (Singh *et al.,*2001), and virtually impossible if stronger race-specific genes are present. In all the above (malting quality, fragrance, QPM, recessive gene, and gene pyramiding) and other similar cases, marker assisted backcrossing is highly justifiable. Once MAB has been completed, it may be continued as Marker Assisted Selection (MAS) within the framework of any breeding method, be it pedigree, recurrent selection, etc. The conditions for marker efficiency will be the same as in MAB, except for the backcross component.

ISSN 2277-9663
AGRES - An International e-Journal

Volume 1 Issue 2 April-June,2012

Gene Synteny

Genes can be discovered using a variety of approaches suggested by Shoemaker *et al.* ,2001. The development of genetic maps in a number of crop species having positional similarity will show the way in better understanding of crop evolution and functioning of genes. This synteny will allow adavances made in one species to have spillover impacts in other species (Gale and Devos, 1998). A comparision of expressed sequence tag (EST) databases from different plants can reveal the diversity in coding sequences between closely and distantly related plants, while mapping of ESTs may elucidate the synteny between those species. For understanding gene functions of a whole organism, functional genomics is now using insertion, mutant isolation, microarrays and proteomics. This information can also used to understand the genetics of metabolic processes, analyze traits controlled by several QTLs, and identify favourable alleles at each locus. The alleles can be combined by simple crossing or using marker assisted selection and / or genetic transformation.

There has been a considerable interest in using synteny to transfer SSR marker isolated from pea, soybean and Medicago. A comparison of linkage maps of Cicer, Pisum, Lens and Vicia has revealed that these legumes share many common linkage groups. (Weeden *et al.,* 1992 and Weeden *et al.,* 2000). The extent of conservation of linkage arrangement may be as much as 40 % of the genome (Weeden *et al.,* 2000). The high level of conservation of linkage groups among Cicer, Pisum, Lens and Vicia suggests that these genera were closely related. There is nearly 60% chance that microsatellite isolated in pea will amplify in chickpea (Edwards *et al.,* 1996), although there is a 20% chance in the reverse direction (Pandian *et al.,* 2000). Based on taxonomic distances, it is expected that a similar trend may be observed between soybean and pigeonpea (Varshney *et al.,* 2012).

Quantitative Trait Loci (QTL)

Plant breeding relies heavily on the science of genetics; the primary goal of a plant breeder is fundamentally different from the primary goal of a geneticist. A plant breeder aims to develop improved cultivars, mainly through selection, whereas a geneticist aims to understand the inheritance and variation of traits. Breeding programs obviously require genetic

ISSN 2277-9663
AGRES - An International e-Journal

Volume 1 Issue 2 April-June,2012

variation for selection to act on, but genetic variation per se is not the main interest of a breeder. Given this context, two general goals of QTL mapping in plants to (i) increase our biological knowledge of the inheritance and genetic architecture (Mackay, 2001) of quantitative traits, both within a species and across related species; and (ii) identify markers that can be used to select for a complex trait. This second goal, which focuses more on breeding than on pure genetics, can be further subdivided into two subgoals:

(a) Identify a few major QTL (i.e., with large estimated effects) that can be introgressed by standard breeding procedures into other germplasm, or

(b) Identify many QTL that can serve as the basis for selection for a complex trait in elite germplasm.

QTL mapping studies have yielded useful biological information in terms of the importance of pleiotropy versus linkage for specific traits (Chung *et al*., 2003) and collinearity in the organization of crop genomes (Gale and Devos, 1998). Furthermore, QTL mapping has served as a springboard for the discovery of the underlying genes through map-based cloning of QTL (Frary *et al.,* 2000), candidate-gene analysis (Pflieger *et al.,* 2001), or comparative mapping (Paterson *et al.,* 1995). Knowledge of the approximate locations of QTL has been used as a starting point for fine mapping by non-QTL mapping approaches or for studying candidate genes that are close to the identified QTL and that may be the actual genes that affect the quantitative trait. At least 20 QTL have been cloned based on their map positions (Price, 2006). If the eventual goal is to clone QTL or identify candidate genes, the penalty of a false positive is severe. The statistical stringency or threshold for declaring the presence of the QTL must therefore be very high. Furthermore, the position of the QTL needs to be mapped precisely relative to closely spaced flanking markers.

Association mapping in plants typically involves finding marker-trait associations among a diverse collection of inbreds with different genetic backgrounds, instead of among recombinant inbreds derived from an F_2 or backcross population between a pair of inbreds as in QTL mapping (Breseghello and Sorrells, 2006). The use of markers that represent polymorphisms at candidate genes would lead to a high resolution in association mapping, although random markers could also be used for genome wide association mapping. Spurious marker-trait associations

ISSN 2277-9663
AGRES - An International e-Journal

Volume 1 Issue 2 April-June,2012

arise due to different genetic backgrounds or pedigrees of the inbreds used, and association mapping needs to account for the population structure among the inbreds that comprise the association-mapping panel (Yu *et al.,* 2006).

Any mapping procedure can detect only those QTL that are polymorphic in the population. The wide assortment of inbreds typically used in association mapping provides the wide genetic diversity needed for discovering a wide array of genes present in the plant species as a whole. This increased genetic diversity, however, often comes at the cost of a decreased mean performance or adaptedness of the germplasm used (Breseghello and Sorrells, 2006). To a geneticist, association mapping is therefore a powerful approach for discovering the genes that underlie quantitative variation (Lazzeroni, 1997). But to a breeder, association mapping with diverse, unadapted germplasm, rather than with elite germplasm, could often represent yet another way to discover additional QTL that would remain largely unexploited in selection for a complex trait, particularly if the contrasting QTL alleles detected by association mapping correspond to mutant forms that have no practical value. These consequences again underscore that the purpose of detecting QTL (e.g., gene discovery versus selection) should therefore be very clearly defined before embarking on a QTL mapping study.

Linkage mapping

The genetic map is defined location/places of specific genetic markers along each linear chromosome of plant species. Genetic mapping is made possible by the fact that the nuclear genome of higher organisms is organized and transmitted as linear units, called chromosomes. Genetic linkage or co-transmission from parent to progeny of genetic markers which are close together on the same chromosome, provides a means for determining the order of DNA markers along the chromosome. The construction of genetic maps is a much more common research approach than it was two decades ago. This increase in mapping activity is, in part, due to technical breakthroughs that have not only made map construction more efficient but have also expended possible uses for genetic maps. A detailed genetic map is a research asset to fungal molecular biologists, pathologists and disease physiologists as well as classical geneticists. Genetic maps spanning part or nearly all of the genomes have now been

ISSN 2277-9663
AGRES - An International e-Journal

Volume 1 Issue 2 April-June,2012

constructed for nearly all important crop plants, Rice, Maize, Cotton, Wheat, Barley, Sorghum, Tomato etc. The process of creating a genetic map has benefits regardless of whether a detailed genetic map is constructed. The most obvious is the identification and characterization of genetic markers.

Physical and Recombination Maps

There are a number of ways that a genetic map, of a whole genome or a specific chromosomal region, can be constructed. The resulting maps can be classified into two major types:

1. Physical maps are the maps where the distance between the markers reflects the actual distance (in bases) between two sites.

2. Recombination maps are the maps where the distance between markers depends on the frequency of genetic recombination between the markers.

Physical Maps

The simplest types of physical maps are those in which markers are assigned to chromosomes. Many organisms, particularly those with small chromosomes, are well suited to this type of physical mapping when RFLP markers are used. Chromosomes can often be separated on agarose gels by various forms of pulse-field gel electrophoresis. The separated chromosomes can than be transferred to blotting membrane and probed with RFLP probes to assign the homologous sequences to the specific chromosomes. This process can be very efficient method of grouping loci on their respective chromosomes, but does not order the loci in the group. Subchromosomal localization (to an arm or chromosome fragment) of RFLP markers may be achieved by electrophoretic separation of cytogenetic stocks such as translocations, or fragmented chromosomes. Deletion stocks are also valuable for mapping DNA markers to chromosomal regions. High resolution physical maps can be generated by probing restriction digested DNA to determine if two probes hybridize to the same fragment. The resolution depends upon the size of the DNA fragments that are generated. Restriction enzymes that cut rarely

ISSN 2277-9663
AGRES - An International e-Journal

Volume 1 Issue 2 April-June,2012

will fragment the chromosomes into only a few large fragments or may not cut some chromosomes at all. This type of map has been generated for the human chromosome arm 23q using the restriction enzyme NotI which cuts very rarely.

Another type of physical map is that where overlapping clones are aligned until a contiguous region of chromosome is cloned. Such contig maps are very cost and labour intensive to construct. Most of the important organisms genomes including human genome are now covered in contigs of cosmid, BAC and YAC clones.

Recombination Maps

To generate a genetic recombination map it is necessary to analyze segregation of markers from individuals from one or more families. The individuals are usually sexual progeny and the genetic markers are segregating as the result of meiosis. It is also possible to generate recombination maps by analysis of mitotic recombinants. This approach may have utility in species in which sexual cycle is absent or difficult to perform.

The most efficient way to create a genetic map by meiotic analysis is to get a sufficient number of markers segregating in a single population. Many meioses (50-200) must be sampled to drive accurate estimates of map distances and to detect linkages between loosely linked (>20 cM) loci. The most efficient analysis, therefore, are those where individual progeny are all or nearly all derived from independent meioses. Recombination frequencies between genes can be calculated directly as the ratio of recombinants/total progeny. Data analysis is therefore straight forward and several mapping programmes are available. When codominant markers are used, segregating diploid populations can give more linkage information per individual progeny than segregating haploid populations. This is particularly true of classic F_2 type of analysis, where F_1s derived from a cross of two homozygous individuals.

The choice of parents is one of the most important considerations in starting a mapping project. An obvious criterion in choosing parents is that any traits that are of prime interest to the investigator must be segregating if they are to be incorporated into the genetic map. Second obvious criterion is that the two parents should have sufficient genetic

distance between them so that they frequently carry different alleles of the molecular markers to be used e. g. two species. In highly polymorphic species, any two genotypes that are not closely related may be sufficient. In species or populations with less polymorphism, this is perhaps the most important factor in determining the efficiency of a mapping project. For example, if one is using RFLPs as markers, and only 10 % of the probes are polymorphic in a population, one would have to screen nearly 10 times as many probes as in a population where nearly all of the probes detected polymorphisms. Genetically distant parents might be selected based on geographical differences, differences in morphological characteristics and parental pedigree distances. The most reliable method, however, is to prescreen potential parents with a subset of markers to be used, and estimate genetic distance directly.

Most genetic mapping populations in plants have been derived from crosses between largely homozygous parents. Backcross populations, F_2 populations, recombinant inbred populations (RI) and doubled haploids (DH) are commonly used in plants. Different objectives can be expedited by the choice of different genetic mapping populations. For initial construction of a primary genetic linkage map in an organism which has not been previously studied, the strong linkage disequilibrium of the F_2 or backcross population permit one to detect linkage between widely-scattered markers

REFERENCES

Adams M. D., Kelley J. M., Gocayne J. D. Dubrick, M. Polymeropoulos, M. H., Xiao H., Merril C. R., Wu A., Olde B., Moreno R. F., Kerlavage A. R., McCombie W. R. and Venter J.C. 1991. Complementary DNA sequencing: expressed sequence tags and human genome project. *Sci*., 252:1651–1656.

Akkaya M. S., Bhagwat A. A. and Cregan P. B. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. *Genet*., 132:1131-1139.

Akopyanz N., Bukanov N. O., Westblom T. U. and Berg D. E. 1992. PCR-based RFLP analysis of DNA sequence diversity in the gastric pathogen *Helicobacter pylori*. *Nucleic Acid Res*., 20:6221-6225.

Beckmann J. S., and Soller M. 1986. Restriction fragment length polymorphisms and genetic improvement of agricultural species. *Euphytica*, 35:111–124.

Breseghello F. and Sorrells M. E. 2006. Association analysis as a strategy for improvement of quantitative traits in plants. *Crop Sci*., 46:1323–1330.

Caetano-Anolles G., Bassam B. J. and Gresshoff P. M. 1991. DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. *Biotech*., 9: 553-

Chung J., Babka H. L., Graef G. L., Staswick P. E., Lee D. J., Cregan P. B., Shoemake, R. C. and Specht J. E. 2003. The seed protein, oil, and yield QTL on soybean linkage group I. *Crop Sci.,* 43:1053–1067.

Edwards K. J., Barker J. H. A., Daly A., Jones C. And Karp A. 1996. Microsatellite libraries enriched for microsatellite sequences in plants. *Biotechniques*, 20: 758-760.

Gale M. D. and Devos K. M. 1998. Comparative genetics in the grasses. *Proc. Natl. Acad. Sci.,* USA 95:1971–1974.

Holland J. B. 2007. Genetic architecture of complex traits in plants. *Curr. Opin. Plant Biol*., 10:156–161.

Huang N., Angeles E. R., Domingo J., Magpantay G., Singh S., Zhang G., Kumaravadivel N., Bennett J. and Khush G. S. 1997. Pyramiding of bacterial blight resistance genes in rice: markers assisted selection using RFLP and PCR. *Theor. Appl. Genet*., 95: 313-320.

Jaccoud D., Peng K., Feinstein D. and Kilian A. 2001. Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Res*., 29:e25.

Jiang G. H., Xu C. G., Tu J. M., Li X. H., He Y. Q. and Zhang Q. F. 2004. pyramiding of insect- and disease-resistance genes into an elite indica, cytoplasm male sterile restorer line of rice, '*Minghui 63'*. *Plant Breed*., 123: 112-116.

Jordan S. A. and Humphries P. 1994. Single nucleotide polymorphism in exon 2 of the BCP gene on 7q31-q35. *Hum. Mol. Genet*., 3:1915.

Jourdren C., Barret P., Horvais R., Foisset N., Delourme R. and Renard M. 1996. Identification of RAPD markers linked to the loci controlling erucic acid level in rapeseed. *Mol. Breed*., 2: 61-71.

King R. C. and Stansfield W. D. 1990. A dictionary of genetics. 4th ed., Oxford University Press, New York-Oxford, pp 188.

Lazzeroni L.C. 1997. Empirical linkage-disequilibrium mapping. Am. J. Hum. *Genet*.,. 62:159–170.

Mackay T. F. C. 2001. The genetic architecture of quantitative traits. *Annual Rev. Genet*., 35:303–339.

Markert C. L. and Moller F. 1959. Multiple forms of enzymes: Tissue, ontogenetic and species specific patterns. *Proc. Natl. Acad. Sci*. USA 45:753-763.

Michelmore R. W., Paran I. and Kesseli R. V. 1991. 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. *Proc. Natl. Acad. Sci.*, USA 88: 9828-9832.3

Pandian A., Ford R. and Taylor P. W. J. 2000. Transferability of sequences tagged microsatellite site (STMS) primers across four major pulses. *Plant Mol. Bio. Reporter*, 18(4): 395a-395h.

Paterson A. H., Lin Y. R., Li Z., Shertz K. F., Doebley J. F., Pinson S. R. M., Liu S. C., Stansel J. W. and Irvine J. E.. 1995. Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. *Sci*., 269:1714–1718.

Pflieger S., Lefebvre V. and Causse M. 2001. The candidate gene approach in plant genetics: A review. *Mol. Breed*., 7:275–291.

Price A. H. 2006. Believe it or not, QTLs are accurate! *Trends Plant Sci*., 11:213–216.

Samis K., Bowley S. and McKersie B. 2002. Pyramiding Mn superoxide dismutase transgenes to improve persistence and biomassproduction in alfalfa. *J. Exp. Bot*., 53: 1343-1350.

Sax K. 1932. The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genet*., 8:552-560.

Semagn K., Bjørnstad A. and Ndjiondjop M. N. 2006b. An overview of molecular marker methods for plants. *African J. Biotech*., 5 (25): 2540-2568.

Shoemaker D. D., Schadt E. E., Armour Y. D., He P., Garett-Engel P. D. and McDonald P. M. 2001. Experimental annotation of the human genome using microarray technology. *Nature*, 409: 922-927.

Singh S., Sidhu J.S., Huang N., Vikal Y., Li Z., Brar D. S., Dhaliwal H. S. and Khush G. S. 2001. Pyramiding three bacterial blight resistance genes (xa5, xa13 and Xa21) using marker-assisted selection into indica rice cultivar PR106. *Theor. Appl. Genet*., 102: 1011-1015.

The Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature, 408:796-815.

The Rice Genome Sequencing Project. 2005. The map-based sequence of the rice genome. *Nature*, 436:793-800.

Toroser D., Thormann C. E., Osborn T. C. and Mithen R. 1995 RFLP mapping of quantitative trait loci controlling seed aliphatic-glucosinolate content in oilseed rape (*Brassica napus* L.). *Theor. Appl. Genet.,* 91: 802-808.

Varshney R. K., Chen Wenbin., Yupeng Li., Bharti,, A. K., Saxena R. K. and Schlueter J. A. 2012. Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. *Nature Biotech*., 30(1):84-92.

Vos P., Hogers R., Bleeker M., Reijans M., Van De Lee T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M. and Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res*., 23: 4407-4414.

Weeden N. F., Ellis T. H. N., Timmerman Vaughan G. M., Simon C. J., Torres A. M., Wolko and Knight R. 2000. How similar are the genomes of the cool season food legumes ? PP. 397-410. In: Linking Research and marketing Opportunities for Pulses in the 21st Century: *Proceedings of the Third International Food Legumes Research conference (Ed., Knight, R.), 22-26 September, 1997. Adelaide, Australia. Kluwer Academic Publishess, London, UK.*

Weeden N. F., Muehlbauer F. J. And Ladizinsky G. 1992. Extensive conservation of linkage relationship between pea and lentil genetic maps. *J. Heredity*, 83: 123-129.

Welsh J. and McClelland M. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res*., 18:7213-7218.

Wexelsen H. 1933. Linkage between quantitative and qualitative characters in barley. *Hereditas* (Lund), 17:323-341.

Williams J. G. K., Kublelik A. R., Livak K. J., Rafalski J. A. and Tingey S. V. 1990. DNA polymorphism's amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.*, 18:6531-6535.

Yoshimura S., Yoshimura A., Iwata A., McCouch S. R., Abenes M.L., Baraoidan M. R., Mew T.W. and Nelson R. J. 1995. Tagging and combining bacterial blight resistance genes in rice using RAPD and RFLP markers. *Mol. Breed*., 1: 375-387.

Yu J., Pressoir G., Briggs W.H., Vroh Bi I., Yamasaki M., Doebley J. F., McMullen M. D., Gaut B. S., Nielsen D., Holland J. B., Kresovich S. and Buckler E. S. 2006. A unified mixedmodel method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet*., 38:203–208.

Zietkiewicz E., Rafalski A. and Labuda D. 1994. Genome fingerprinting by simple sequence repeats (SSR)-anchored PCR amplification. *Genomics*, 20:176-183.