Journal of Bio-Molecular Sciences (JBMS) (2014) 2(3): 78-85.

Journal of Bio-Molecular Sciences (JBMS)
ISSN: 2311-4630
www.directsciences.com/jbms

Molecular Markers in Plant Genotyping

Rehan Naeem

Department of Biotechnology and Genetic Engineering, Kohat University of Science and Technology (KUST), Pakistan

Received 10 Oct. 2014; Accepted 20 Oct. 2014; available online 30 Oct. 2014

Abstract: Morphological characterization is the first step in description and classification of the germplasm. But biochemical markers like isozymes, resistance to chemicals and seed storage proteins are also being used for genotyping studies in plant breeding. Now days, molecular markers are the markers of attention for advanced research. These can be classified into two categories including non-Polymerase Chain Reaction (PCR) based techniques and PCR-based techniques. Several new array chip techniques that utilize DNA hybridization coupled with labeled nucleotides and Next-Generation sequencing techniques have also been developed. The techniques that are promising for plant genotyping are the Restriction fragment length polymorphism (RFLP), Amplified Fragment Length Polymorphism (AFLP), Random Amplified Polymorphic DNA (RAPD), Microsatellites or Simple Sequence Repeat (SSR), and Single Nucleotide Polymorphism (SNP). The choice of a molecular marker technique depends upon the reliability, ease of analysis and level of polymorphisms. The merger of genetics and genomics has unraveled the new insights by defining various metabolic, regulatory and developmental pathways, but rigorous investigations still need to be completed.

Key words: Markers, genotyping, plants

Introduction

Genetic improvement in crop plants is mainly focused on the selection of optimal germplasm having desired traits that include higher yield, improved nutritional quality, appropriate color or fragrance etc. These traits are not measured directly from the plants but they are evaluated from some sources that are closely associated with the trait of interest. Such tags which used to select the superior germplasm are called as markers. These markers are useful in an array of plant breeding and genetics and can be classified into two categories including classical markers and DNA markers (Xu, 2010). Classical markers include morphological markers, cytological markers and biochemical markers. Morphological characterization is the first step in description and classification of the germplasm. For this purpose, various numerical taxonomic techniques have been successfully utilized to classify and measure pattern of genetic diversity in germplasm. Cytological markers involve the visualization of structural features of chromosomes that have been shown by chromosome karyotype and bands. Biochemical markers like isozymes,

*Correspondence to: Rehan Naeem, Department of Biotechnology and Genetic Engineering, Kohat University of Science and Technology (KUST), Pakistan; E-mail: gulerehan@gmail.com

resistance to chemicals and seed storage proteins are also used for genetic diversity studies (Jiang, 2013).

DNA marker is defined as a particular segment of DNA that differs among individuals at the nucleotide level. These markers may or may not correlate with phenotypic expression of a trait and offer numerous advantages over conventional morphological markers and isoenzymes. DNA markers have developed into many systems based on different polymorphism-detecting techniques or methods including southern blotting, nuclear acid hybridization, PCR – Polymerase Chain Reaction and DNA sequencing (Collard et al., 2005), such as Restriction fragment length polymorphism (RFLP), Amplified fragment length polymorphism (AFLP), Random Amplified Polymorphic DNA (RAPD), Simple Sequence Repeats (SSR), Single Nucleotide Polymorphism (SNP), etc. An ideal DNA marker should be polymorphic and evenly distributed throughout the genome. These markers can provide adequate resolution of genetic differences and must be simple, quick, inexpensive having linkages to distinct phenotypes. They are stable and detectable in all tissues regardless of growth, differentiation, development and status of the cell. They are also not confounded by the environment, pleiotropic and epistatic effects. These markers can be classified into two categories including non-PCR based techniques or hybridization based techniques and PCR-based techniques. A nuclear acid hybridization based Southern blotting, (Southern, 1975), and PCR, a Polymerase Chain Reaction (Mullis, 1990) both require electrophoresis as detection systems that include PAGE - Polyacrylamide Gel Electrophoresis, AGE - Agarose Gel Electrophoresis, CE - Capillary Electrophoresis. The variation in DNA samples for a specific region can be identified by the product features, such as band size and mobility. In addition to Sothern blotting and PCR, more detection systems have been also developed. For instance, several new array chip techniques that utilize DNA hybridization coupled with labeled nucleotides and Next-Gen Sequencing techniques have also been developed. Among the techniques that have been extensively used and are particularly promising for application to plant genotyping are the Restriction fragment length polymorphism (RFLP), Amplified Fragment Length Polymorphism (AFLP), Random Amplified Polymorphic DNA (RAPD), Microsatellites or Simple Sequence Repeat (SSR), and Single Nucleotide Polymorphism (SNP). These markers are grouped for the discriminative power of determining individuals as homozygosous (dominant marker) or heterozygous (co-dominant marker) (Hartl, 1988). Their principle, mode of inheritance, advantages, disadvantages and degree of polymorphism are shown in Table 1.

Amplified Fragment Length Polymorphism (AFLP)

As a nuclear hybridization marker, AFLP combines the power of RFLP with the flexibility of PCR-based technology by ligating primer recognition sequences (adaptors) to the restricted DNA and selective PCR amplification of restriction fragments using a limited set of primers (Vos et al., 1995). The primer pairs used for AFLP usually produce 50–100 bands per assay. AFLP involves restriction digestion of genomic DNA (about 500 ng) with two restriction enzymes, a rare cutter (6-bp

Table 1. Principal, mode of inheritance, advantages and limitations of different DNA markers used in plant genotyping

Marker class	Principle	Mode of inheritance	Advantages	Disadvantages	Degree of Polymorphism
RAPD	Differences in primer annealing sites	Dominance	Quick, simple and inexpensive	Poor reproducibility	Low – medium
AFLP	Differences in the presence or absence of recognition site and differences in the primer annealing sites	Dominance	Multiple loci	Large amount of DNA required	Low – medium
SNP	Difference in the sequences at single-nucleotide level	Dominance	Extremely degraded DNA samples can be used	Each marker has less alleles	Medium – high
RFLP	Presence / Absence of recognition sites	Co-Dominance	Reliable	Time consuming, laborious	Low – medium
SSR	Differences in number of repeats of microsatellite motifs	Co-Dominance	Technically simple, robust and reliable	Time and cost intensive initial establishment	Medium – high

*Source: (Boopathi, 2013)

recognition site, *Eco*RI, *Pst*I or *Hind*III) and a frequent cutter (4- bp recognition site, *Mse*I or *Taq*I). The adaptors are then ligated to both ends of the fragments to provide known sequences for PCR amplification. The double-stranded oligonucleotide adaptors are designed in such a way that the initial restriction site is not restored after ligation. Therefore, only the fragments which have been cut by the frequent cutter and rare cutter will be amplified. This property of AFLP makes it very reliable, robust and immune to small variations in PCR amplification parameters and it also can produce a high marker density. The AFLP products can be separated in high-resolution electrophoresis systems. The fragments in gel-based or capillary DNA sequencers can be detected by dye-labeling primers radioactively or fluorescently. The number of bands produced can be manipulated by the number of selective nucleotides and the nucleotide motifs used.

AFLP can also be used to distinguish closely related individuals at the subspecies level and map genes. AFLP are regularly employed in plant mapping including establishing linkage groups and assessing the degree of relatedness or variability among cultivars. For high-throughput screening approach, fluorescence tagged primers are also used for AFLP analysis.

The amplified fragments are detected on denaturing polyacrylamide gels using an automated ALF-DNA sequencer with the fragment option (Huang and Sun, 1999).

Randon Amplified Polymorphic DNA (RAPD)

RAPD is a dominant marker that revealed differential PCR amplification of genomic DNA using short random oligonucleotide sequences (mostly ten bases long) that generate banding pattern, which is dependent upon the length and size of both the primer and the target genome. Usually differential banding pattern generated is due to the rearrangements or deletions between oligonucleotide primer binding sites in the genome (Williams et al., 1991). The PCR products (up to 3 kb) are separated by agarose gel electrophoresis and imaged by ethidium bromide (EB) staining. It offers no prior knowledge of the genome and can be employed across species using universal primers. RAPD is predominantly dominant markers and informative markers that yield high levels of polymorphism. In RAPD, no marker development is required, and the primers are non- species specific and can be universal. It offers many advantages such as higher frequency of polymorphism, rapidity, simplicity, use of fluorescence and requirement of a few nanograms of DNA. RAPD markers have been used extensively in crop improvement, identification of somatic hybrids, evaluation and conservation of genetic resources, DNA profiling, population genetics and gene mapping. However, RAPD is limited in its application because of its inconsistency in PCR reactions as annealing temperature; template DNA concentration and Mg 2+ ion concentration are sensitive for RAPD that affect the banding pattern. Arbitrarily Primed Polymerase Chain Reaction (AP-PCR) and DNA Amplification Fingerprinting (DAF) are independently developed methodologies, which are variants of RAPD. For AP-PCR (Welsh and McClelland, 1990), a single primer (about 10–15 nucleotides long) is used. The technique involves amplification for initial two PCR cycles at low stringency. Thereafter, the remaining cycles are carried out at higher stringency by increasing the annealing temperature. RAPD markers are used to develop other types of PCR-based markers, such as Sequence Characterized Amplified Region (SCAR), Single Nucleotide Polymorphism (SNP), etc. SCARs are identified by PCR amplification using sequence- specific oligonucleotide primers (Paran and Michelmore, 1993). Development of SCARs involves cloning the amplified products of arbitrary marker techniques and then sequencing the two ends of the cloned products. The sequence is thereafter used to design specifiic primer pairs of 15-30 bp which amplify single major bands of the size similar to that of cloned fragment. Polymorphism is either retained as the presence or absence of amplification of the bands. Co-dominant SCARs are more informative for genetic mapping than dominant arbitrary-primed molecular markers, as they can be used to screen pooled genomic libraries by PCR and for physical mapping, defining locus specificity as well as comparative mapping and homology studies among related plant species. Thus, SCARs have higher reproducibility; however cloning and sequencing are still laborious in SCAR development. To avoid this problem, extended random primer amplified region (ERPAR) has been developed (Wang et al., 2000). Similar to SCAR, an ERPAR uses specific primer pairs derived from RAPD primers by adding bases sequentially to their 3’ ends (Boopathi, 2013).

Restriction Fragment Length Polymorphism (RFLP)

RFLP are the first generation of DNA markers and has been utilized extensively in human genetics, animal genetics, plant genetics and germplasm characterization. In RFLP, DNA polymorphism is detected by hybridising a chemically labelled DNA probe to a Southern blot of sample DNA which has digested with restriction endonucleases (Botstein et al., 1980). Thus, RFLP generates differential banding profile which is generated due to nucleotide substitutions or DNA rearrangements like insertion or deletion or single-nucleotide polymorphisms in recognition site of the restriction enzymes. The DNA probe used for the detection of polymorphism is a radioactively labelled DNA sequence that hybridises with one or more fragments of the restriction enzyme digested DNA sample after they have separated by gel electrophoresis. Thus, RFLP results in unique banding pattern characteristic to a specific genotype at a specific locus. These are co-dominant, locus-specific, polymorphic and are highly reproducible. However, RFLP is not widely used in linkage mapping since it is time consuming, involves expensive and radioactive/toxic reagents and requires large quantity of high-quality genomic DNA. The requirement of prior sequence information for probe generation further increases the complexity of the methodology. These limitations led to the develoment of a new set of less technically complex methods that are based on PCR (Boopathi, 2013).

By using RFLP technique, a high-throughput marker has been developed that are cleaved amplified polymorphism sequence (CAPS), also known as PCR-RFLP. RFLP markers are predominantly used in 1980s and 1990s, but since last decade fewer direct uses of RFLP markers in genetic research have been reported because it is considered as laborious and demands too much pure DNA. The CAPS marker utilized DNA sequences of mapped RFLP markers to develop PCR-based markers thereby eliminating the tedious DNA blotting (Komori and Nitta, 2005). The CAPS make out the restriction fragment length polymorphisms caused by single base changes like SNPs, insertions/deletions, which modify restriction endonuclease recognition sites in PCR amplicons. The CAPS assays are performed, by digesting locus-specific PCR amplicons with one or more restriction enzyme, followed by separation of the digested DNA on agarose or polyacrylamide gels. The primers are synthesized based on the sequence information available in databank of genomic or cDNA sequences or cloned RAPD bands. The CAPS markers are co-dominant and locus specific and have been used to distinguish between plants that are homozygous or heterozygous for alleles. The technique is, however, limited by mutations, which create or disrupt a restriction enzyme recognition site (Boopathi, 2013).

Microsatellites (SSRs)

Microsatellites (SSRs) are short tandem repeats (STRs). These are Di-, tri- and tetra-nucleotide repeats widely distributed throughout the genomes of plants and animals. The copy number of these repeats varies among individuals and is a source of polymorphism. Chloroplast microsatellites are particularly effective markers for studying mating systems, gene flow via both pollen and seeds and uniparental lineage. In contrast mt DNA, which typically has a size of 10 MDa per mitochondrial genome, is far more complex and characterized by molecular heterogeneity observed as classes of circular chromosomes that vary in size and relative abundance.

The DNA sequences flanking microsatellite regions are usually conserved.

The unique sequences bordering the SSR motifs provide templates for specific primers to amplify the SSR alleles through PCR. The PCR-amplified products can be separated in high-resolution electrophoresis systems and the bands can be visually recorded by fluorescent labeling or silver-staining. SSR markers are characterized by their hyper-variability, reproducibility, co-dominant nature, locus-specificity, and random genome-wide distribution in most cases (Liu and Cordes, 2004). SSR technique requires nucleotide information for primer design, labor-intensive marker development process and high start-up costs for automated detections. Since the 1990, microsatellites have proven to be an extremely valuable tool for genome mapping in many organisms (Knapik et al., 1998) and have been extensively used in constructing genetic linkage maps, QTL mapping and marker-assisted selection but their applications span over different areas ranging from kinship analysis, to population genetics and conservation/management of biological resources (Jarne and Lagoda, 1996).

Single Nucleotide Polymorphism (SNP)

Single nucleotide polymorphism (SNP) is the most abundant marker system both in animal and plant genomes and has recently emerged as the new generation molecular markers for various applications. Moreover, unlike microsatellites their power comes not from the number of alleles but from the large number of loci that can be assessed (Foster et al., 2010). Variations at single-nucleotide level in genome sequence of individuals of a population are known as SNPs (Jordan and Humphries, 1994). Being binary or co-dominant status, they are able to efficiently discriminate between homozygous and heterozygous alleles. SNPs are present within coding sequences of genes, non-coding regions of genes or in the intergenic regions between genes at different frequencies in different chromosome regions. Improvements in sequencing technology and availability of an increasing number of EST sequences have made direct analysis of genetic variation at the DNA sequence level. High-throughput genotyping methods, including DNA chips, allele-specific multiplex PCR and primer extension approaches, make SNPs especially attractive as genetic markers. Because of these technological improvements, SNPs are highly suitable for automation and are used for construction of ultra-high-density genetic maps. A convenient method for detecting SNPs is RFLP (SNP- RFLP) or by using the CAPS marker technique. However, high costs for marker development and costly equipment make SNP as costly marker. A high- throughput genome analysis method called diversity array technology (DArT), based on microarray platform, has also been developed for the analysis of DNA polymorphism (Jaccoud et al., 2001). As an alternative to DNA markers, RNA have also been used as template to develop special kinds of molecular markers including complementary DNA-AFLP (cDNA-AFLP), cDNA-SSCP and RNA fingerprinting by arbitrarily primed PCR (RAP-PCR).

Conclusion

In general, the choice of a molecular marker technique depends upon the reliability and ease of analysis, statistical power and confidence of revealing polymorphisms. So, genomics has not only brought an innovative level of hope for the development of novel types of markers and unraveling the secrets of complex traits but also provide the possibility for association mapping, QTL Mapping and marker assisted selection. Functional genomics approaches can also be used to generate information about gene function and genetic interactions, not only among and between gene

complexes but also in response to environmental stimuli. It has been proposed that the power available through the merger of genetics and genomics might lead to further unraveling of metabolic, regulatory and developmental pathways, but rigorous investigations still need to be completed.

References

Boopathi, N. M. 2013. Genotyping of Mapping Population. In: Genetic Mapping and Marker Assisted Selection: Basics, Practice and Benefits. Springer, London. pp. 39-80.

Botstein, D., White, R. L., Skolnick, M., Davis, R. W. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32: 314-333.

Collard, B. C. Y., Jahufer, M. Z. Z., Brouwer, J. B., Pang, E. C. K. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. Euphyt. 142: 169-196.

Foster, J. T., Allan, G. J., Chan, A. P., Rabinowicz, P. D., Ravel, J., Jackson, P. J. and Keim, P. 2010. Single nucleotide polymorphisms for assessing genetic diversity in castor bean (*Ricinus communis*). BMC Plant Biol. 10: 13-23.

Hartl, D. L. and Jones, E. W. 2005. DNA Structure and DNA manipulation. In: Genetics: analysis of genes and genomes. 5th ed., Sudbury: Jones and Bartlett Pub. pp. 36-85.

Huang, J., Sun, M. 1999. A modified AFLP with fluorescence labeled primers and automated DNA sequencer detection for efficient fingerprinting analysis in plants. Biotechnol. Tech. 14: 277–278.

Jarne, P. and Lagoda, P. J. L. 1996. Microsatellites, from molecules to populations and back. Trends in Ecol. Evol. 11: 424-429.

Jiang G. L. 2013. Molecular Markers and Marker-Assisted Breeding in Plants. InTech. http://creativecommons.org/licenses/by/3.0.

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

Knapik, E. W., Goodman, A., Ekker, M., Chevrette, M., Delgado, J., Neuhauss, S., Shimoda, N., Driever, W., Fishman, M. C. and Jacob, H. J. 1998. A microsatellite genetic linkage map for zebrafish (Danio rerio). Nat. Genet. 18: 338-343.

Komori, T., Nitta, N. 2005. Utilization of CAPS/dCAPS method to convert rice SNPs into PCR-based markers. Breed. Sci. 55: 93-98.

Liu, Z. J. and Cordes, J. F. 2004. DNA marker technologies and their applications in aquaculture genetics. Aquacul. 238: 1-37.

Mullis, K. 1990. The unusual origin of the polymerase chain reaction. Sci. Am. 262 (4): 56–61.

Paran, I., Michelmore, R. W. 1993. Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theor. Appl. Genet. 85: 985-999.

Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98: 503-517.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van-de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., Zabeau, M. 1995. AFLP: a new technique for DNA fingerprinting. Nucl. Acid. Res. 23: 4407-4414.

Wang, X., Zhiyuan, F., Sanwen, H., Peitian, S., Yumei, L., Limei, Y., Mu, Z., Dongyu, Q. 2000. An extended random primer amplified region (ERPAR) marker linked to a dominant male sterility gene in cabbage (*Brassica oleracea* var capitata). Euphyt. 112: 267-273.

Welsh, J., McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucl. Acid. Res. 18: 7213-7218.

Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A., Tingey, S. V. 1991. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucl. Acid. Res. 18: 6531-6535.

Xu, Y. 2010. Molecular Breeding Tools: Markers and Maps. In: Molecular Plant Breeding. (Ed.): Xu, Y. CAB International, Wallingford, UK. pp. 21-43.