Journal of Agriculture and Crops
ISSN(e): 2412-6381, ISSN(p): 2413-886X
Vol. 4, Issue. 8, pp: 86-92, 2018
URL: http://arpgweb.com/?ic=journal&journal=14&info=aims

Academic Research Publishing Group

Original Research **Open Access**

Application of Molecular Markers SNP and DArT in Plant Breeding: A Review Paper

Tafere Mulualem Emrey
Ethiopian Institute of Agricultural Research-Pawe Agricultural Research Center Pawe, Ethiopia

Abstract

Molecular markers (DNA markers) have entered the scene of genetic improvement in a wide range of horticultural crops. Among the major traits targeted for improvement in horticultural breeding programmes are disease and pest resistance, fruit yield and quality, tree shape, floral morphology, drought tolerance and dormancy. The development of molecular techniques for genetic analysis has led to a great increase in the knowledge of horticultural genetics and understanding and behavior of their genomes. These molecular techniques in particular, molecular markers, have been used to monitor DNA sequence variation in and among the species and create new sources of genetic variation by introducing new and favorable traits from landraces, wild relatives and related species and to fasten the time taken in conventional breeding. Today, markers are also being used for, genetic mapping, gene tagging and gene introgression from exotic and wild species.

Keywords: Breeding; Marker; Molecular; Plant.

1. Introduction

The efficiency of the breeding and selection process can be assessed in many different ways including the ultimate success of the varieties released and the frequency with which new varieties are produced. A major cost and logistical issue in plant breeding are the actual number of lines that need to be carried through the evaluation and selection phases of a program. Molecular markers have proved to be a powerful tool in replacing bioassays and there are now many examples available to show the efficacy of such markers. According to Langridge [1] the use of molecular markers to track loci and genome regions in crop plants is now routinely applied in many breeding programs. The location of major loci is now known for many disease resistance genes, tolerances to abiotic stresses and quality traits. Improvements in marker screening techniques have also been important in facilitating the tracking of genes. For markers to be effective, they must be closely linked to the target locus and be able to detect polymorphisms in material likely to be used in a breeding program.

The focus of this paper is to show two technologies that appear to be best suited for whole-genome profiling i.e. SNP and DArT. SNP markers are expected to substantially replace other molecular marker technologies that developed over the last quarter of the 20th century such as RFLP [2]. The most popular SNP typing assays and the most promising SNP technology platforms are briefly reviewed here. The slow uptake of SNP markers in plants, especially those plants for which limited resources are invested, is caused partly by the small or nonexistent market demand and partly by the lack of sequence information. DArT is an upcoming technology that avoids most of the problems of lower-throughput typing technologies and does not require sequence information. Here I describe the principles of DArT and the current status of development of this technology including examples of DArT application to crop species. I conclude with a comparison of the advantages and disadvantages of two technologies in the context of plant breeding.

Diversity arrays technology (DArT) is a novel genotyping technique that was originally developed for rice [3] but has been applied to cassava, wheat .etc

2. What is Marker?

Molecular marker is a sequence of DNA or protein that can be screened to reveal key attributes of its state or composition and thus used to reveal genetic variation. Many of these markers mentioned above require the use of gel electrophoresis, are time consuming and expensive. The emerging new molecular markers such as Single Nucleotide Polymorphism (SNPs) techniques do not always need these electrophoresis-based assays. SNPs are excellent markers for association mapping of genes controlling complex traits and provide the highest map resolution [4, 5] [6]. SNPs are robust in usage and polymorphisms are identifiable and there are several methods that can be used to detect them. SNPs are the most frequent type of variation found in DNA [5, 7] and their discovery together with insertions/deletions has formed the basis of most differences between alleles. SNPs can thus be explained as any polymorphism between two genomes that is based on a single nucleotide exchange. In plants, studies on the occurrence and nature of SNPs are beginning to receive considerable attention, particularly in *Arabidopsis* where over 37, 000 SNPs have been identified through the comparison of two accessions [8].

The SNP marker may be considered the ultimate genetic marker as they represent the finest resolution of a DNA sequence and SNP marker is the most frequently found DNA sequence variations the development of high-throughput methods for the detection of SNPs has led to a revolution in their use as molecular markers within DNA

sequence data. To increase the importance of SNP markers, various databases have been constructed [9]. It has been reported in maize that there occurs a frequency of one non-coding SNP per 31 bp and 1 coding SNP per 124 bp in 18 maize genes assayed in 36 inbred lines [10].

3. The Application of Molecular Marker in Breeding

Recent developments that have occurred in molecular markers for many crop species have major implications for the future of the technology. There are three key components that are particularly significant. First, for many species, we now have markers closely linked to many traits of importance in the breeding programs. Indeed, for major crop species, we have markers for more loci that can be screened in a conventional breeding program. Second, we have tools that allow marker scanning of the whole genome. Of particular importance has been the development of microsatellite or SSR markers that now form the basis for analysis and allow highly multiplexed SSR screens. This trend will continue as newer, cheaper marker screening based on SNPs become available. The technological advances have improved our capacity for whole genome screens. Third, through association mapping projects we have, or are in the process of developing, whole genome fingerprints for many key lines and varieties of importance in breeding programs.

The new marker systems have several important implications for the future of marker-assisted selection (MAS) and breeding strategies in general. Existing strategies for MAS were initiated with a view of markers as providing a rapid and cheap alternative to bioassays and they have largely been used in this role. While highly successful, this strategy does not fully exploit the technology. The key limitation to an expansion of the scale and complexity of marker use is the size of the populations that would be required if one were to try and select for alleles at a large number of loci simultaneously. A further important feature of recent advances has been related to how we best take advantage of the genome information that has been generated for major crop species.

Molecular markers have been taken, in recent years, to refer to assays that allow the detection of specific sequence differences between two or more individuals. However, it should be recognized that isoenzyme and other protein-based marker systems also represent molecular markers and were in wide use long before DNA markers became popular. One of the earliest types of DNA-based molecular markers, restriction fragment length polymorphisms (RFLPs), was based around the detection of variation in restriction fragment length detected by Southern hybridisation. The types of sequence variation detected by this procedure could be caused by single base changes that led to the creation or removal of a restriction endonuclease recognition site or through insertions or deletions of sufficient size to lead to a detectable shift in fragment size. This technique has been largely superceded by microsatellite or simple-sequence repeat (SSR) markers and is now rarely used in screening material for breeding programs, but it remains an important research tool. SSR markers detect variation in the number of short repeat sequences, usually two or three base repeats. The number of such repeat units has been found to change at a high frequency and allows the detection of multiple alleles.

The large expansion of DNA, particularly EST, sequence databases has now opened the opportunity for the identification of single nucleotide polymorphisms, SNPs. These occur at varying frequencies depending on the species and genome region being considered. In Arabidopsis SNP frequencies of 0.007–0.0104 have been measured [11, 12] while in maize a range of 0.00047–0.0037 has been measured [13, 14]. SNPs are widely seen as providing the key advantage of multiple detection systems many of which, such as mass spectroscopy, offer high throughput at low detection cost. Importantly, new array based screening methods, such as DArT [3] appear to offer still cheaper assays due to their very high multiplexing capability. Interestingly, molecular markers may be coming full circle with protein markers again being proposed as viable genetic markers for MAS. Mass spectrometric methods for mass fingerprinting of proteins and for the analysis of low molecular weight proteins, again opens the option for high throughput protein screening. In these cases, single amino acid changes in protein sequence can often be detected and this provides a means for revealing variation in the corresponding DNA coding sequence. According to Theo, *et al.* [15], the potential for applying molecular markers is clear, but there are also some pitfalls that need to be mentioned. These issues can be divided in two categories: technological and economical. The technological issue mainly is related to the fact that most commonly used markers are considered to be selectively neutral, i.e., they are not subject to selection pressures. The second issue is of economic nature. Molecular marker analysis generally is expensive. Obviously, all resources spend on this technology can not be spend on other genebank activities.

4. Types of Marker

Generally the marker can be classified as: i) Morphological, ii) Biochemical and iii) Molecular. One marker differs from the other by genomic abundance, polymorphism level, locus specificity, reproducibility, technical requirements and financial investment. A Single Nucleotide Polymorphism (SNP) is a DNA sequence variation occurring when a single nucleotide in the genome differs between members of a species. According to Register, these variations in DNA sequence, called *polymorphisms*, can be associated or linked with different forms (alleles) of nearby genes involved with particular traits. The polymorphism, or difference, is the clue researchers need to find the gene of interest. Allele Specific Hybridization (ASH) focuses on the most common form of genetic variation in plants: single nucleotide polymorphism (SNP). Because SNPs occur frequently throughout the plant genome, they offer enormous potential in the discovery and detection of important genes in crops. Just as the name indicates, SNPs are identified by a single nucleotide base change in the genetic code at a specific location on the chromosome. On discovered, these single-base differences can be used as molecular markers for genes of interest.

A series of different molecular marker systems, which became available during the last two decades, can be broadly classified into three classes: (i) the first generation molecular markers, including RFLPs, RAPDs and their modifications; (ii) the second generation molecular markers, including SSRs, AFLPs and their modified forms, and (iii) the third generation molecular markers including ESTs and SNPs. Since ESTs are mainly used for studies on functional genomics, SNPs are the only new generation molecular markers for individual genotyping needed for molecular marker-assisted selection (MAS). In view of a causal similarity of SNPs with some of these marker systems and the fundamental difference with several other marker systems, the molecular markers have also been classified into SNPs (due to sequence variation, e.g. RFLP) and non-SNPs (due to length variation, e.g. SSRs).

During the last few years, SNPs have also attracted the attention of plant molecular biologists, and one may hope that in future SNP maps will be prepared and used extensively in many plant systems. In plant systems, the SNPs seem to be more abundant than even those in the human genome, so that in preliminary studies conducted in wheat, one SNP per 20 bp, and in the maize genome, one SNP per 70 bp have been recorded in certain regions of these genomes. One can only hope that SNPs will be developed expeditiously in all major crops in a large scale and will be extensively utilized in future for a variety of crop improvement programmes, although non-availability of adequate sequence data may limit this activity.

5. Single Nucleotide Polymorphism/SNP/

Single nucleotide variations in genome sequence of individuals of a population are known as SNPs. SNPs are the most abundant molecular markers in the genome. They are widely dispersed throughout genomes with a variable distribution among species. The SNPs are usually more prevalent in the non-coding regions of the genome. Within the coding regions, when an SNP is present, it can generate either non-synonymous mutations that result in an amino acid sequence change, or synonymous mutations that not alter the amino acid sequence. Synonymous changes can, however, modify mRNA splicing, resulting in phenotypic differences. Improvements in sequencing technology and an increase in the availability of the increasing number of EST sequences have made analysis of genetic variation possible directly at the DNA level.

The majority of SNP genotyping analyses are based on: allele-specific hybridization, oligonucleotide ligation, primer extension or invasive cleavage. Genotyping methods, including DNA chips, allele-specific PCR and primer extension approaches based on SNPs, are particularly attractive for their high data throughput and for their suitability for automation. They are used for a wide range of purposes, including rapid identification of crop cultivars and construction of ultra high-density genetic maps.

5.1. Definition and Importance in Selection

As suggested by the acronym, an SNP (single nucleotide polymorphism) marker is just a single base change in a DNA sequence, with a usual alternative of two possible nucleotides at a given position. For such a base position with sequence alternatives in genomic DNA to be considered as an SNP, it is considered that the least frequent allele should have a frequency of 1% or greater. Although in principle, at each position of a sequence stretch, any of the four possible nucleotide bases can be present, SNPs are usually biallelic in practice. One of the reasons for this, is the low frequency of single nucleotide substitutions at the origin of SNPs,

What is the reason for the increasing popularity of SNPs, whereas in terms of genetic information provided, as simple bi-allelic co-dominant markers, they can be considered as a step backwards when compared to the highly informative multi-allelic microsatellites? In fact, the more recent SNP concept has basically arisen from the recent need for very high densities of genetic markers for the studies of multifactorial diseases, and the recent progress in polymorphism detection and genotyping techniques. In recent years, scientists have begun to recognize the promise SNPs offer for targeting important genes. Because they represent the most common type of genetic variation in plants, SNPs can have significant effects on resistance to disease, performance under adverse environmental conditions and other economically important traits. “SNPs are particularly valuable because they offer the potential of ultrahigh throughput and highly automated analysis,” says Dr. Jim Register, research coordinator of analytical nucleic acid technologies at Pioneer. “SNP analysis is a ‘yes’ or ‘no’ proposition (Is the sequence of interest there or not?). SSR marker measures a varying number of sequence repeats.”

Although numerous approaches for SNP discovery have been described, including some also currently used for genotyping, the main ones are based on the comparison of locus-specific sequences, generated from different chromosomes. The simplest, when targeting a demand region for instance containing candidate genes, is to perform direct sequencing of genomic PCR products obtained in different individuals. However, on a large scale, this approach tends to be costly due to the need for locus-specific primers, is limited to regions for which sequence data is available, and produces a diploid sequence in which it is not always easy to distinguish between sequencing artifacts and polymorphism when double peaks, as expected in heterozygotes, are observed.

5.2. Use of Snps in Crop Improvement Programmes

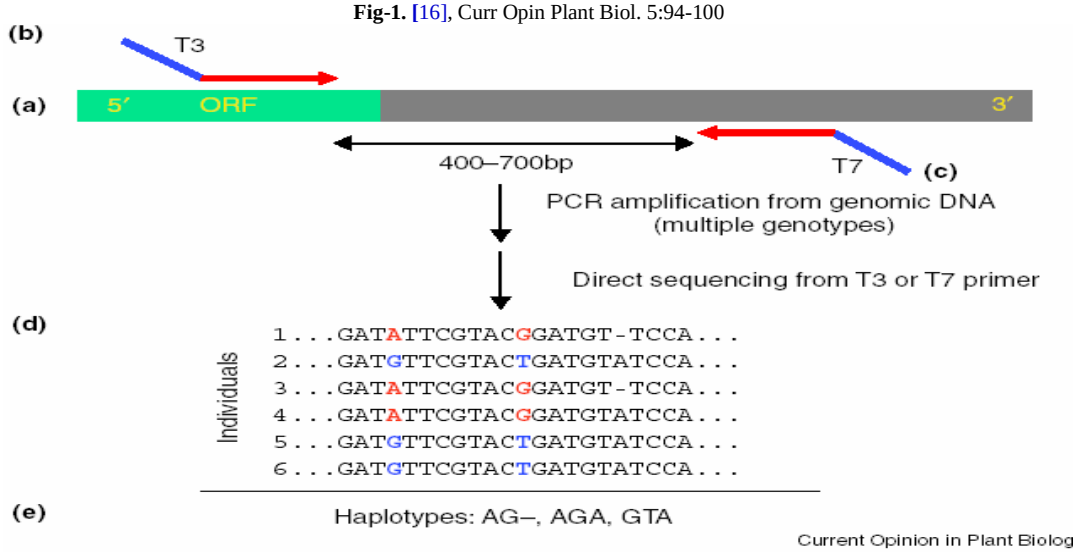
SNP can be found within a gene or may be found in its close proximity. When found within a gene, it may or may not be responsible for the mutant phenotype, but in either case, it can be used for positional cloning of the gene in question. Once a large collection of SNPs is available, their use will depend on whether or not genetic determinants for all traits are included in those SNPs. For instance in humans, according to *causal hypothesis* the final collection of SNPs will include all genetic determinants, so that the allelic associations among SNPs will be irrelevant. In contrast to this, according to *proximity hypothesis*, most genetic determinants of diseases will not be

included even in a sample of several hundred thousand SNPs, so that allelic association among SNPs can be very important and will be used for positional cloning. We believe that in plants, most SNPs will not be genetic determinants and therefore associations among SNPs and the traits of economic value will be of major interest to the plant breeders. SNP (Single Nucleotide Polymorphisms) relatively large availability of DNA sequence information enables to identify a large number of SNP which are the most abundant type of variation among individuals of the same species.

5.3. Methods Used For Discovery and Identification of New Snps

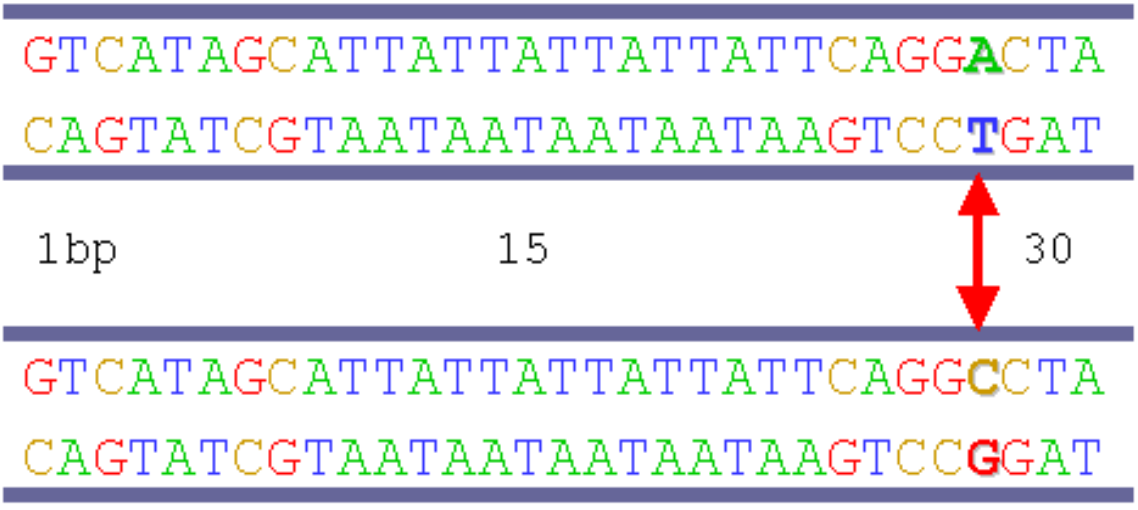
RFLPs, RAPDs and SSRs which were the markers of choice during the last two decades, need gel-based assays and are, therefore, time consuming and expensive. Therefore, emphasis is now shifting towards the development of molecular markers, which can be detected through non gel-based assays. One of the most popular of these non gel-based marker systems is SNP, which represents sites, where DNA sequence differs by a single base. This polymorphism has been shown to be the most abundant, so that at least one million SNPs should be available, only in the non-repetitive transcribed region of the human genome.

Fig-1. [16], Curr Opin Plant Biol. 5:94-100



A newer type of marker that has now been made available through new sequencing technologies is single nucleotide polymorphisms (SNPs). These polymorphisms are single-base substitutions between sequences. SNPs occur more frequently than any other type of marker, and are very near to or even within the gene of interest.

Fig-2. Interpreting SNPs: The schematic drawing of a single nucleotide polymorphism shows two DNA fragments (top and bottom) sharing the same sequence for 31 base pairs, except one. In position 28, an A-T (top) has changed to a C-G (bottom)



6. Diversity Arrays Technology/Dart/

DArT is a generic and cost-effective genotyping technology. It was developed to overcome some of the limitations of other molecular marker technologies such as RFLP, AFLP and SSR. DArT is an alternative method to time-consuming hybridisation-based techniques, typing simultaneously several thousand loci in a single assay. DArT is particularly suitable for genotyping polyploid species with large genomes, such as wheat. This technology generates whole-genome fingerprints by scoring the presence/absence of DNA fragments in genomic representations generated from samples of genomic DNA. DArT technology consists of several steps: (i) complexity reduction of

DNA; (ii) library creation; (iii) the microarray of libraries onto glass slides; (iv) hybridisation of fluoro-labelled DNA onto slides; (v) scanning of slides for hybridisation signal and (vi) data extraction and analysis. DArT acts by reducing the complexity of a DNA sample to obtain a "representation" of that sample. The main method of complexity reduction used relies on a combination of restriction enzyme digestion and adapter ligation, followed by amplification. However, an infinite range of alternative methods can be used to prepare genomic representations for DArT analysis. DArT markers for a new species are discovered by screening a library of several thousand fragments from a genomic representation prepared from a pool of DNA samples that encompass the diversity of the species. The microarray platform makes the discovery process efficient because all markers on a particular DArT array are scored simultaneously. For each complexity reduction method, an independent collection of DArT markers can be assembled on a separate DArT array. The number of markers for a given species, therefore, is only dependent on: (i) the level of genetic variation within the species (or gene pool); and (ii) the number of complexity reduction methods screened.

6.1. Definition and Importance in Selection

DArT was developed to provide a practical and cost-effective whole-genome fingerprinting tool [3]. DArT has three key attributes of interest to plant breeders and scientists studying and managing genetic diversity: (a) it is independent from DNA sequence, (b) the genetic scope of analysis is defined by the user and easily expandable, and (c) the method provides for high throughput and low-cost data production [17]. A DArT marker is a segment of genomic DNA, *the presence of which* is polymorphic in a defined *genomic representation*. DArT markers are biallelic and behave in a dominant (present vs absent) or co-dominant (2 doses vs 1 dose vs absent) manner. Genetic diversity is the raw material available to plant breeders. By productively recombining genetic diversity, plant breeders have been successfully producing, year after year, improved cultivars of the major domesticated species used in the world's diverse agricultural systems. Molecular genetic markers offer a powerful tool to accelerate and refine this process. Existing genetic marker (genotyping) technologies, mostly developed for applications in human health, have also been applied successfully to agricultural species, but their cost remains prohibitive for most agricultural applications. This is particularly true for species for which no molecular data and very limited resources are available. Because of the limitations of existing marker technologies, we have developed Diversity Arrays Technology (DArT), a novel method to discover and score genetic polymorphic markers. DArT is a sequence-independent, high-throughput method, able to discover hundreds of markers in a single experiment. DArT markers are typed in parallel, using high-throughput platforms, with a low cost per data point. With DArT, plant breeders, plant ecologists, as well as the managers of the germplasm collections, will be able to perform genetic analysis in a cost-effective and high-throughput manner. DArT fingerprints will be useful for accelerating plant breeding, and for the characterization and management of genetic diversity in domesticated species as well as in their wild relatives. We have developed DArT successfully for rice, barley, wheat, and cassava. We have also produced a dedicated data management and analysis package, a key part of the technology, entirely built from Open Source components. Work is in progress to establish DArT for other species of importance to tropical agriculture: pigeon pea, sorghum, and chickpea.

6.2. Identification of Polymorphic Dart Markers

To identify the polymorphic markers, a *complexity reduction method* is applied on the metagenome, a pool of genomes representing the germplasm of interest. The genomic representation obtained from this pool is then cloned and individual inserts are arrayed on a microarray resulting in a "discovery array." Labeled genomic representations prepared from the individual genomes included in the pool are hybridized to the discovery array. Polymorphic clones (DArT markers) show variable hybridization signal intensities for different individuals. These clones are subsequently assembled into a "genotyping array" for routine genotyping. The discovery of polymorphic DArT markers and their scoring in subsequent analysis does not require any DNA sequence data. This makes the method applicable to all species, regardless of how much DNA sequence information is available for that species. However, DArT markers are sequence-ready clones of genomic DNA.

6.3. Applications of DArT Markers

DArT markers can be used as any other genetic marker. With DArT, comprehensive genome profiles are becoming affordable for virtually any crop, regardless of the level of molecular information available for the crop. We anticipate that DArT genome profiles will be used for the recognition and management of biodiversity, for example in germplasm collections. Identification of duplicate accessions and a better understanding of the genetic relationships between the accessions could help to control the costs of maintaining these collections.

In plant breeding, DArT genome profiles will enable breeders to map QTL in one week, thereby allowing them to focus on the most crucial factor in plant breeding: reliable and precise phenotyping. Once many genomic regions of interest are identified in many different lines, DArT profiles accelerate the introgression of a selected genomic region into an elite genetic background (for example by Marker Assisted Back Crossing). Furthermore DArT profiles can be used to guide the assembly of many different regions into improved varieties. For that purpose, dense genome cover is essential in order to follow many regions simultaneously. Because of the large number of lines to be typed, high throughput and affordability are critical factors in this context.

Table-1. Ongoing DArT projects in different species

Species	No. representations tested	No. clones assayed
Rice	14	26,112
Barley	10	21,504
Wheat	5	14,592
Apple	3	1,920
Cassava	4	9,216
Perennial rye grass	5	5,376
Pigeon pea	4	5,376
Sorghum	2	1,536
Fungal pathogens of barley	4	5,376
Arabidopsis	1	1,536
Mouse	2	1,536
Bovine	2	1,536
Sheep	5	3,840

7. Summary and Conclusion

Molecular markers are used also to assess plant response to climate change, which is a major issue at a global level. Changes, such as rapid warming, have been seen to cause a decrease in the variability of those loci controlling physical responses to climate. Jump and Peñuelas conducted a review of climatic factors correlated with microgeographical genetic differences, and the various molecular markers used for each study. They concluded that although phenotypic plasticity buffers against environmental changes over a plant's life cycle, it will weaken over time as climatic event become more extreme and over longer time spans. The assessment and maintenance of genetic diversity, through the use of molecular markers is crucial as it provides a repository of adaptability to environmental and other changes.

Molecular markers are useful for a variety of purposes relevant to crop improvement. The most important of these uses is the indirect marker assisted selection (MAS) exercised during plant breeding. For this purpose, molecular markers need to be amenable to automation and high throughput approaches. However, the gel-based assays that are needed for most molecular markers are time consuming and expensive, limiting their utility. The new generation molecular markers, called single nucleotide polymorphisms (SNPs) do not always need these gel-based assays. They are also the most abundant of all marker systems known so far, both in animal and plant genomes. A large number of SNPs have already been developed in the human genome, some of them proving useful for diagnosis of diseases. A beginning has also been made in the development and use of SNPs in higher plants, including some crop and tree species.

Hopefully in future, they will be used in plants in a big way. Several approaches can be used for discovery of new SNPs and about a dozen different methods are now available for SNP genotyping. Some of these methods are also suitable for automation and high throughput approaches. These methods, in principle, make a distinction between a perfect match and a mismatch (at the SNP site) between a probe of known sequence and the target DNA containing the SNP site. The target DNA in most of these methods is a PCR product, except in some cases like 'invasive cleavage assay', and 'reduced representation shotgun (RRS)' devised and used recently. The different methods of SNP discovery and detection, along with examples of some known uses of SNPs in plant systems are described in this article.

During the last 3–5 years, SNPs have emerged as the new generation molecular markers, which have already been developed in large number for the human genome. In the next year, while many more human SNPs will be developed and mapped, SNPs will also be produced in several crops. In parallel with their discovery and development, SNPs are already being used in humans for detection of association with a variety of diseases. In crop plants, however, only a beginning has been made in the area of SNP discovery and detection. In future, they will certainly be used in a number of crops, not only for studies involving associations with a number of traits of economic value, but also for the study of genetic diversity and variety identification. Enormous genomic and cDNA sequence data that are accumulating in the databases will be extremely useful in future for discovery of new SNPs. A number of gel-based and non gel-based methods will also be used for detection of already characterized SNPs and for genotyping of populations at these SNP sites. Newer methods will also be developed for this purpose. This will be facilitated due to automation and high throughput approaches that are already available for work on SNP. In DArT, several hundred polymorphic markers are identified in parallel. The efficiency of this marker discovery effort is only dependent on the level of genetic diversity within the species. For example, 5-10% of wheat and barley DArT clones and 25-30% of cassava DArT clones were polymorphic. The same platform is used for both discovery and scoring of markers, therefore no assay development, apart from consolidating all polymorphic markers into a single genotyping array, is required after the marker discovery.

References

[1] Langridge, J. P., 2004. "Correlating instrumental and sensory analyses of flavour." PhD thesis, University of Nottingham.

[2] Andrzej, K., Eric, H., Peter, W., and Damian, J., 2005. "The fast and the cheap: SNP and DArT-based whole genome profiling for crop improvement." Available: www.researchgate.net/publication/228755323

[3] Jaccoud, D., Peng, K., Feinstein, D., and Kilian, A., 2001. "Diversity arrays: A solid state technology for sequence information independent genotyping." *Nucleic Acids Research*, vol. 29, p. e25.

[4] Botstein, D. and Risch, N., 2003. "Discovering genotypes underlying human phenotypes: past successes for Mendelian disease, future approaches for complex disease." *Nat. Genet*, vol. 33, pp. 228-237.

[5] Brookes, A. J., 1999. "The essence of SNPs." *Gene*, vol. 234, pp. 177-186.

[6] Bhattramakki, D., Dolan, M., Hanafey, M., Wineland, R., Vaske, D., Register, J. C., Tingey, S. V., and Rafalski, A., 2002. "Insertion-deletion polymorphisms in 3' regions of maize genes occur frequently and can be used as highly informative genetic markers." *Plant Mol. Biol.*, vol. 48, pp. 539-547.

[7] Cho, R. J., Mindrinos, M., Richards, D. R., Sapolsky, R. J., Anderson, M., Drenkard, E., Dewdney, J., Reuber, T. L., Stammers, M., *et al.*, 1999. "Genome-wide mapping with biallelic markers in Arabidopsis thaliana." *Nat. Genet*, vol. 23, pp. 203-207.

[8] Jander, G., Norris, S. R., Rounsley, S. D., Bush, D. F., Levin, I. M., and Last, R. L., 2002. "Arabidopsis map-based cloning in the post-genome era." *Plant Physiol.*, vol. 129, pp. 440-450.

[9] ChangKug, K., UngHan, Y., GangSeob, L., SungHan, P., Young-Joo, S., HwanKi, L., and JangHo, H., 2009. "An integrated database to enhance the identification of SNP markers for rice (Oryza sativa) and Chinese cabbage (Brassica rapa)." *Afr. J. Biotechnol.*, vol. 8, pp. 5253-5259.

[10] Ching, A., Caldwell, K. S., Jung, M., Dolan, M., Smith, O. S., Tingey, S., Morgante, M., and Rafalski, A. J., 2002. "SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines." *BMC. Genet.*, vol. 3, p. 19.

[11] Kawabe, A., Hideki, I., fel Ryohei, T., and Naohiko, T. M., 1997. "Nucleotide Polymorphism in the Acidic Chitinase Locus (ChiA) Region of the Wild Plant Arabidopsis thaliana." *Mol. Biol. Evol.*, vol. 14, pp. 1303-1315.

[12] Purugganan, M. D. and Suddith, J. I., 1998. "Molecular population genetics of the Arabidopsis CAULIFLOWER regulatory gene: Nonneutral evolution and naturally occurring variation in floral homeotic function." *Proc. Natl. Acad. Sci. USA.*, vol. 95, pp. 8130–8134.

[13] Hilton, H. and Gaut, B. S., 1998. "Speciation and domestication in maize and its wild relatives: evidence from the globuline-1 gene." *Genetics*, vol. 150, pp. 863–872.

[14] Wang, D. G., Fan, J. B., Siao, C. J., Berno, A., Young, P., and Sapolsky, R., 1998. "Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome." *Science*, vol. 280, pp. 1077-1082.

[15] Theo, J. L., Van, H., and Rob, V. T., 2002. "Molecular markers: Tools to improve gene bank efficiency."

[16] Rafalski, A., 2002. "Applications of single nucleotide polymorphisms in crop genetics." *Curr. Opin. Plant Biol.*, vol. 5, pp. 94-100.

[17] Eric, H., Peter, W., Mona, A., and Vanessa, C., 2005. "Diversity arrays technology: A novel tool for harnessing the genetic potential of orphan crops." Available: www.researchgate.net/publication/265072468