Indian Journal of Agricultural Sciences **89** (11): 1755–63, November 2019/Review Article
https://doi.org/10.56093/ijas.v89i11.95287

Molecular markers for characterization and conservation of plant genetic resources

AEJAZ AHMAD DAR[1], REETIKA MAHAJAN[2] and SUSHEEL SHARMA[3]

Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu, Chatha, Jammu 180 009, India

Received: 28 July 2018; Accepted: 07 March 2019

ABSTRACT

For properly understanding the molecular basis of phenotypic variation, accelerating breeding programmes and exploiting genetic resources of a particular crop, there is a need to explore the genome of the crop. Earlier conventional methods (phenotyping) were used by the breeders for identification and conservation of genetic resource. The conventional method is laborious for assignment of known and unknown plant taxa. Thus, to overcome the limitation of the conventional method, scientists developed molecular marker systems for conservation of genetic resources. The main aim of using molecular markers is to detect genetic variation in DNA sequence within and among the species. Molecular markers also find their role in producing new sources of genetic variation by introduction of advanced and useful traits from related species. From last two decades, different molecular markers were used for genetic analysis of crops. These markers include hybridization based markers (RFLP), PCR based markers (RAPD, AFLP, SSR, ISSR, SCAR, SRAP, SCoT) and sequence based markers (SNP, DArT, NGS). Besides many genes in chloroplast, mitochondria and ribosomes act as molecular markers that are useful for the phylogenetic studies and DNA barcoding of plants. This review provides an overview of the different molecular systems and their utilization in improving crop yield and productivity and conserving genetic resources.

Key words: AFLP, Conservation, DNA barcoding, Genetic Resources, Molecular markers, NGS, Phylogentic study, SSR

Modern agriculture mainly deals with the conservation and utilisation of valuable genetic resources. Given that biodiversity is fundamental to the breeders' programmes, management of genetic resources via exploration, collection, characterization, evaluation, and conservation should be highlighted. In addition, the danger of genetic erosion in wild species and primitive forms of crop plants and the associated likely consequences for agriculture delineated the potential direction for future classical approaches to be further promoted by the molecular techniques (Arzani and Ashraf 2017). The need for correct identification applies to cultivars and accessions, independently of their mode of conservation. The identification and characterisation are sometimes difficult by conventional methods. Since morphological markers are environment dependent and time consuming, molecular approaches are implemented presently in cultivar identification and breeding programmes. Molecular markers have several advantages in differentiating errors, cultivar identification and stability under different environmental conditions and further, it accelerates the breeding process by selecting elite cultivars by tracking certain genes or genotypes among crossed populations (Dar *et al.* 2017a, Sharma *et al.* 2017).

DNA markers act as strong tools for the estimation and evaluation of genetic variability, DNA fingerprinting, phylogenetic relationships within and among cultivars, and increased selection efficiency through marker-assisted selection for complex traits in plant breeding (Altintas *et al.* 2008). DNA markers have a great respect to plant breeders for precisely mapping different interacting genes that are linked to desirable agronomic traits such as abiotic stress tolerance, pest resistance, disease resistance and food quality. Unlike biochemical and morphological markers, molecular markers are environmentally stable, provide a significantly high genetic polymorphism, and at the same time permit analysis at any developmental stage (Dar *et al.* 2014). The limitations associated with morphological markers for analysing genetic diversity in different plant species are largely removed by the usage of DNA markers. All these DNA markers have technical differences in terms of speed, cost, degrees of polymorphism, amount of DNA required and accuracy of genetic distance. Different molecular techniques exploit nucleotide polymorphisms that arise from different types of mutations such as insertions or deletions,

Present address: [1]National Post Doctorate Fellow (aejazdbt.pbiot@gmail.com), [2]Senior Research Fellow (pisces418@gmail.com), [3]Assistant Professor (drsusheelsharma@rediffmail.com), Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu, Chatha, Jammu, India.

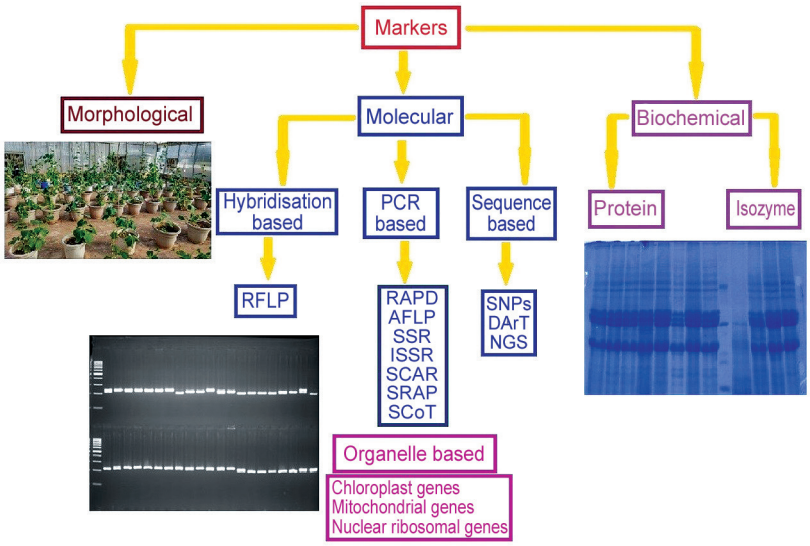

Fig 1 A systematic overview of different types of molecular markers.

point mutations, or error in replications of tandem-repeats (Smykal *et al.* 2012).

Molecular markers

The molecular markers are DNA sequences at specific regions of the genome inherited in a Mendelian fashion which either encode or not necessarily any traits but are unaffected by the environment. For a marker to be ideal and good, it should give sufficient information per run, inexpensive, easy to perform and score, repeatable, error free or with low error rate, identify many alleles (Kumar *et al.* 2009). Based on the methodology for the detection of markers, molecular markers are classified into three groups. First group includes hybridization markers such as restriction fragment length polymorphism (RFLP), second group includes PCR based markers such as random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), simple sequence repeats or microsatellites (SSR), inter-simple sequence repeats (ISSR), sequence characterized amplified region (SCAR), sequence-related amplified polymorphism (SRAP) and the third group includes DNA sequencing based molecular markers such as single nucleotide polymorphism (SNP), diverse array technology (DArT) and next generation sequencing (NGS) (Fig 1). A brief comparison of the advantages and limitations of different marker system is given in Table 1.

Hybridization based markers

Restriction fragment length polymorphism: RFLP technique was devised by the English scientist Alec Jeffreys during his research on hereditary diseases. Although RFLPs play a pivotal role in many plant breeding applications, their main limitation lies in the identification of genotypes with low relationship and the complexity of the technique (Wunsch and Hormaza 2002). RFLP markers were used to fingerprint the *Pennisetum purpureum* germplasm to determine the genetic diversity, heterozygosity and hybridization. Linkage analyses revealed 64 markers linked to QTL of 26 plant traits. These QTL-linked markers formed the basis for genetic study of important biomass productivity and quality traits of the plant (Smith *et al.* 1993). Paull *et al.* (1998) studied associations in wheat between agronomically significant traits and RFLP phenotypes. The comparisons of the RFLP phenotypes containing disease resistance genes of stem-rust, leaf-rust and yellow-rust enabled identification of specific allele's characteristic of these regions. Hence, RFLP can be useful for characterisation of wheat germplasm and identification of chromosome segments associated with agronomic traits.

PCR based markers

Randomly amplified polymorphic DNA: RAPD was the first and the simplest of all the PCR based molecular markers developed for assessing genetic diversity between plant species. RAPD analysis was done on 84 F_2 plants through bulked segregant analysis (BSA) approach for identifying markers linked to plant height trait. Out of 200 RAPD primers, only two markers (OPF04700 and OPA091375) were found to be linked to the plant height gene as they were present only in the open-tall plants, while absent in compact dwarf plants (Dhanasekar *et al.* 2010). Dar *et al.* (2017a) compared genetic diversity in *Sesamum indicum* L. with RAPD and SSR markers. They reported maximum PIC with SSRs (0.194) compared to RAPDs (0.186). Higher marker index was observed with RAPDs (1.426) than with SSRs (0.621). The RAPD primer RPI-B11 and SSR primer S16 were the most informative in terms of describing genetic diversity. Based on seed coat colour, white and brown seeded varieties were clustered with each other, while black seeded varieties remained distanced from the cluster. Genetic diversity of 54 Indian pea genotypes was studied by using 30 RAPD primers (Thakur *et al.* 2018). Genetic similarity estimates based on Jaccard's similarity coefficient ranged from 0.34 (Kaza-2/IC-218991) to 0.89 (IC-209118/IC-209123) with an average similarity index of 0.54 exhibiting considerable diversity among the pea genotypes. Such diversity can potentially be used in pea breeding programmes to broaden the genetic base of existing varieties.

Amplified fragment length polymorphisms: AFLP was invented by the Dutch company, Key gene, in order to overcome the non reproducible results produced with RAPD. AFLP is fastest, inexpensive, and most reproducible technology for estimating genetic diversity and high-density genetic maps of desirable traits (Chial 2008). Obert *et al.* (2000) screened 36 downy mildew resistant alfalfa plants and 36 susceptible plants for DNA polymorphisms using 14 AFLP primer combinations. Four AFLP fragment markers, ACACTC208, ACACTC150, ACACAT216 and ACACTC486 were found to be significantly associated

Table 1 Advantages and limitations of commonly used marker systems (Sarwat 2012, Robarts and Wolfe 2014, Yu *et al.* 2018)

Marker	Advantages	Limitations
RFLP	-Co-dominant -Genomic abundance high -Highly reproducible -Better genome exposure -Applicable across the species -No need for sequence information -Reliably used in plants	-Need high-quality DNA -Laborious -Complex to automate -Radioactive labeling
RAPD	-Genomic abundance high -Better genome coverage -Sequence information not required -Perfect for automation -Requires less DNA -No radioactive labeling	-No need of probe information -Dominant markers -Not reproducible -Not well tested
ISSR	-Sequence not required -Reliable, fast and cost effective -Less quantity of DNA required -Distinguish closely related species	-Non homology of similar sized fragements -Non-reproducible -Primers less specific
SSR	-Easy to automate -Genomic abundance high -Highly reproducible -High polymorphism -Multiple alleles -Moderately genome coverage -No radioactive labeling	-Not well-examined -Cannot suitable across species -Sequence information needed
Minisat-ellites	-Highly polymorphic -Multiallelic -High reproducibility -Low cost	-Many informative bands per reaction -Band profiles cannot be interpreted in terms of loci and alleles
AFLP	-High polymorphism -High genomic abundance -Can be used across species -No need for sequence information -Useful in preparing counting maps -Works with smaller RFLP fragments	-Very tricky due to changes in materials used -Not reproducible -Very good primers needed
SCAR	-Highly reproducible -Quick and simple -Locus specific	-Need gene sequence to design markers -Sometimes radioactive isotopes is required
SRAP	-Simple and efficient -Highly reproducible and versatile -Reasonable throughput rate -Targets ORFs for amplification	-Limitations not yet reported

with disease susceptibility or resistance. Two of the AFLP markers, ACACTC208 and ACACTC486 were significantly associated with resistance in the F_1 and S_1 progeny. These identified markers can be used for the improvement of crop for disease resistance. Lu *et al.* (2008) mapped QTLs affecting root morphological traits in *Brassica rapa* by AFLP and RAPD markers. They identified 18 QTLs for the 3 root traits, 7 QTLs for taproot thickness, 5 QTLs for taproot length, and 6 QTLs for taproot weight. The QTLs detected for root morphological traits, may provide a basis for marker-assisted selection to improve agricultural productivity in root-crop breeding. El-Esawi *et al.* (2018) evaluated the variation and relationships of 25 French pea accessions using fatty acid profiling and AFLP markers. The percentage of crude oil varied from 2.6–3.5%, with a mean of 3.04%, and the gene diversity of the AFLP markers varied from 0.21–0.58, with a mean of 0.41. Three AFLP markers (E-AAC/M-CAA, E-AAC/M-CAC, and E-ACA/M-CAG) were found significantly associated with the crude oil content trait. A correlation was also revealed between the genetic diversity level and salt tolerance degree in pea genotypes. The results of this study might be used to improve pea crop.

Simple sequence repeats: SSRs repeats or microsatellites are the most variable tandemly repeated DNA sequences (Table 2). Major molecular markers generated by microsatellite sequences are: SSLPs (Simple Sequence Length Polymorphism), STMSs (Sequence Tagged Microsatellite Site), SCARs (Sequence Characterized Amplified Region), SNPs (Single Nucleotide Polymorphisms), and CAPS (Cleaved Amplified Polymorphic Sequences). Nirmala *et al.* (2016) confirmed the purity of hybrids between salt resistant and susceptible genotypes of mungbean using SSR markers. Micro-satellite markers may assist breeding by early selection of the genotypes compared to the phenotypic screening. SSRs were also used in estimating genetic diversity and population structure in 104 genotypes of cucumber using 23 SSR primers. The information obtained will aid researchers for identification of genotypes, molecular breeding, and further improvement of crop (Dar *et al.* 2017b). A collection of 96 common bean genotypes were characterized using 25 candidate gene based SSR markers associated with yield traits. High level of genetic diversity observed could be used for genetic improvement of common bean targeting yield attributing traits (Gupta *et al.* 2018).

Expressed Sequence Tags (ESTs) are single-read sequences that are produced from partial sequencing of mRNA pool and reverse transcribed into cDNA. Twenty eight EST-SSRs were selected from NCBI database for determining genetic diversity among 27 castor accessions. Genetic similarity coefficient and polymorphic information content ranged from 0.24–0.83 and 0.28–0.49, respectively. These results indicated a medium level of diversity present in castor genotypes that can be useful for plant breeders to plan efficient hybrid breeding programme (Thatikunta *et al.* 2016). An initiative was taken by Sharma *et al.* (2019) to characterise genetic diversity of 40 pea genotypes using 24 EST-SSRs. The PIC value ranged from 0.095–0.500 at an

Table 2 Classification of SSR (microsatellites)

Classification	Type of SSRs	References
Based on number of nucleotides/repeat	i. Mononucleotide (A)n ii. Dinucleotide (CA)n iii. Trinucleotide (CGT)n iv. Tetranucleotide (CAGA)n v. Pentanucleotide (AAATT)n vi. Hexanucleotide (CTTTAA)n (n = number of variables)	Kalia *et al.* 2011
Based on the arrangement of nucleotides in the repeat motifs	i. Pure or perfect or simple perfect (CA)n ii. Simple imperfect (AAC n ACT (AAC)n + 1 iii. Compound or simple compound (CA)n (GA)n iv. Interrupted or imperfect or compound imperfect (CCA)n TT (CGA)n + 1	Wang *et al.* 2009
Based on location of SSR in genome	Nuclear (nuSSRs) Chloroplastic (cpSSRs) Mitochondrial (mtSSRs)	Kalia *et al.* 2011

average of 0.349. Jaccard's similarity coefficient ranged from 0.17 (Azad Pea1/ IC-219010) to 0.92 (IC-219010/219008) with an average value of 0.45 indicating considerable genetic diversity. Bayesian model-based programme divided the accessions into two major populations. The information gathered from EST-SSR markers may assist breeders to use breeding programmes for the improvement of target traits.

Inter-simple sequence repeat: The ISSRs are semi-arbitrary markers that are amplified with a primer sequence complementary to a targeted microsatellite sequence. Choudhary *et al.* (2013) investigated genetic diversity of 38 accessions of *Cicer arietinum*, 6 of *C. reticulatum*, and 4 of *C. echinospermum* by 31 ISSR markers. The study revealed moderate level of diversity within accessions with an average PIC value of 0.125. The cluster and the principal coordinate analysis revealed a clear differentiation between wild and cultivated accessions. Marker-trait associations in 20 wheat varieties were examined under drought conditions using phenotypic and ISSR markers (Motawea *et al.* 2015). Polymorphism information content values ranged from 0.08–0.28. Plant height and 1000-kernel weight showed significant regressions under normal irrigation conditions. The study demonstrated that ISSR markers are promising tools for estimating genetic diversity, conservation, identification, marker-assisted selection and utilization of germplasm for crop improvement through conventional and molecular breeding approaches. Su *et al.* (2017) assessed the genetic diversity of 102 genotypes of a novel oil crop *C. brevistyla* by ISSR markers. Cluster analysis identified 7 groups of populations, and analysis of molecular variance indicated that the variation within populations (75%) was predominantly greater than variations among counties (18%) and among regions (7%). The highly differentiated populations of *Camellia brevistyla* can serve as a platform for the oil industry as a germplasm resource.

Sequence characterized amplified regions: SCARs amplify DNA by using specific pair of oligonucleotide primers designed from cloned RAPD fragments linked to a trait of interest. The SCARs are highly reproducible and locus specific, and are normally involved in marker assisted selection and gene mapping studies (Yu *et al.* 2018). Basha and Sujatha (2007) identified population specific SCAR markers in *Jatropa curcas* L. The SCAR primer pair ISPJ1 amplified a 543 bp in all the Indian populations, whereas ISPJ2 amplified a 1,096 bp specific to the Mexican genotype. Similarly, Suprasanna *et al.* (2008) characterized radiation induced and tissue culture derived dwarf banana plants by using a SCAR marker. The amplification with the SCAR primer produced an expected fragment of 1500 bp in normal but not in the dwarf types. Thus, SCAR markers serve well in identification of specific plant populations, gene tagging, and marker assisted selection and gene mapping studies. Binyamin *et al.* (2015) identified resistant or tolerant genotypes of mungbean using SCAR markers linked to yellow mosaic virus resistant genes. The SCAR markers amplified the desired band only in the resistant and tolerant genotypes, whereas no amplification was found in susceptible genotypes. SCAR markers can be, thus, useful in future breeding and varietal developmental programs.

Sequence related amplified polymorphism: SRAP is a PCR based dominant marker system used to amplify DNA with primers targeting their open reading frames. The SRAPs is an efficient and simple marker for producing genome-wide fragments of high versatility and reproducibility (Robarts and Wolfe 2014). SRAP markers along with morphological traits were used to estimate genetic relationship among 40 pea varieties (Esposito *et al.* 2007). Relationships among varieties revealed by SRAP markers were significantly correlated with the agronomic traits, suggesting that the two systems give similar estimates of genetic relations among the pea varieties. Ahmed *et al.* (2013) identified 5 SRAP markers linked to leaf chlorophyll content, leaf senescence and cell membrane stability in wheat under water-stressed condition. Consequently, these SRAP markers can be useful in molecular breeding programmes for the improvement of agricultural crops and their productivity. Guenni *et al.* (2016) analyzed 43 accessions of *Pistacia vera* using SRAP markers. A high genetic differentiation and high gene flow was reported among groups. Analysis of molecular variance revealed that 73.88% of the total genetic diversity was present within groups, whereas the remaining 26.12% occurred among groups. The information obtained will be helpful in conservation strategies and improvement programs of *Pistacia* spp.

Start codon targeted (SCoT) polymorphisms: It is a novel and recent DNA marker based on the short conserved region flanking the ATG start codon in plant genes (Collard and Mackill 2009). SCoT markers along with ISSRs were

used for identification and genetic comparison of 23 Chinese mango accessions (Luo *et al.* 2011). The cultivars were clustered into two major groups and were in accordance with their known origins and phenotypic characteristics. The study indicated that SCoT markers better represented the actual relationships among mango accessions than ISSR markers. The SCoT marker system is, thus, useful for identification and genetic diversity analysis of mango cultivars. Que *et al.* (2014) assessed the genetic diversity among 107 sugarcane accessions using 20 SCoT markers. PIC values ranged from 0.783–0.907 with an average value of 0.861, suggesting abundant genetic diversity among sugarcane cultivars. The knowledge of genetic diversity obtained can provide a foundation for managing sugarcane germplasm, constructing core collection and regional variety distribution. Yang *et al.* (2015) studied genetic diversity and taxonomic status of 26 date plum persimmon accessions based on morphological and SCoT markers. Significant variability was observed in 21 morphological characters. Genetic similarity of accessions ranged from 0.42–0.83 indicating higher genetic variation.

Organelle based markers

Chloroplast genes: Due to the availability of complete chloroplast genome sequences in many genera, it is convenient to access the variability of chloroplast regions at low taxonomic levels. Identification of variable loci in chloroplast genomes will be useful for molecular systematics and DNA barcoding (Daniell *et al.* 2016). The chloroplast genes that are used currently for phylogenetic studies at different taxonomic levels are *atpB*, *atpB-rbcL*, *matK*, *ndhF*, *rbcL*, *rpl16*, *rps4-trnS*, *rps16*, *trnH-psbA*, *trnL-F*, *trnS-G* etc (Gao *et al.* 2008). Many chloroplast genes such as *atpF-H*, *matK*, *psbK-1*, *rbcL, rpoB*, *rpoC1* and *trnH-psbA* act as candidate markers used for plant DNA barcoding (Kress *et al.* 2005). Garris *et al.* (2005) analyzed 234 rice accessions with 169 nuclear SSRs and two chloroplast loci. Genetic structure revealed 5 distinct groups corresponding to *indica*, *aus*, *aromatic*, *temperate japonica*, and *tropical japonica* rices. A closer evolutionary relationship was observed between the *indica* and the *aus* and among the *tropical japonica*, *temperate japonica*, and *aromatic* groups. The study can be useful in developing a population-based framework for analyzing the molecular diversity in *O. sativa*. Dong *et al.* (2012) studied the chloroplast genomes of 12 genera for searching highly variable regions. Out of the 23 loci found, the most variable were intergenic regions such as*ycf1-a, trnK, rpl32-trnL* and *trnH-psbA*followed by $trnS^{UGA}$-$trnG^{UCC}$, *petA-psbJ, rps16-trnQ, ndhC-trnV, ycf1-b, ndhF, rpoB-trnC, psbE-petL,* and *rbcL-accD.*

Mitochondrial genes (mtDNA): The mtDNA separated by large regions of noncoding DNA became popular at species level in phylogenetics and population genetic studies. The cytochrome oxidase (CO) of electron transport chain is slowly evolving gene as compared to other mitochondrial genes and is widely used for constructing phylogenetic tree. COI and COII genes were used for species and population analyses of parasitoids and COI gene was well established for barcoding insects and animals (Hollingsworth *et al.* 2011). Cytochrome-b gene is also the most useful marker in recovering phylogenetic relationships among closely related taxa. It has been reported that mitochondrial 12S rRNA gene was used for identification and phylogenetic analysis of the Indian leopard's species and wildlife forensic biology (Pandey *et al.* 2007). Shu *et al.* (2016) estimated the diversity of 39 broccoli accessions, including 19 cytoplasmic male sterility lines (CMS) and 20 hybrids using mitochondrial markers. The 39 accessions were classified into five groups based on six *orf*138-related and two SSR markers. As many as eight molecular markers were identified to detect CMS types during broccoli breeding. The information gained from such study will be important in future studies on the origins and molecular mechanisms of CMS in broccoli.

Nuclear ribosomal genes: Ribosomal RNA is considered to be the best marker for studying phylogenetic relationship because of the presence of conserved and variable domains, and universal nature. It was reported that the conservation of 16S rRNA gene among *Bacillus* species was used to differentiate between different bacterial groups (Clarridge 2004). 5S rRNA is also highly conserved and a rapidly evolving molecule used for phylogentic analysis. Similarly, 28S rRNA became available recently for inferring phylogentic relationship among major metazoan groups. The internal transcribed spacer (ITS) region in the ribosomal cistron is a primary fungal barcode used for identification of inter- and intraspecific variation in broadest range of fungi. The nuclear ribosomal large subunit unlike small subunit in fungi is a popular phylogenetic marker that had superior species resolution than some other taxonomic groups, such as ascomycete yeasts (Schoch *et al.* 2012). Benor (2018) studied phylogenetic analysis of the 144 *Corchorus* accessions using nuclear ribosomal DNA internal transcribed spacer (ITS) region. The majority of African species formed a distinct clade separated from the other pantropically distributed species. Several other endemic species from Australia, New Caledonia, and tropical America were found within this distinct clade, indicating dispersal from Africa to the rest of the pantropics. Further phylogenetic analysis with additional molecular and morphological markers would help to find out the biogeography and evolution of the genus *Corchorus*.

DNA sequencing based markers

Single nucleotide polymorphism: Single nucleotide polymorphism is the most wonderful molecular marker found frequently in the genome showing a variable distribution among species. SNPs are the DNA markers of choice as they are rapidly replacing SSRs for their role in plant breeding and genetics (Edwards and Batley 2010). Lee *et al.* (2017) developed SNP markers from the chloroplast and nuclear genomic sequences for identifying specific Korean specific ecotypes of *C. tricuspidata* via amplification refractory mutation system (ARMS)-PCR

and high resolution melting (HRM) curve analyses. Wang *et al.* (2011) established a first-generation linkage map of *Jatropa* using a mapping panel containing two backcross populations with 93 progeny. They mapped 506 markers (216 microsatellites and 290 SNPs from ESTs) onto 11 linkage groups. The total map length was 1440.9 cM with an average marker space of 2.8 cM. After blasting 222 *Jatropha* ESTs against EST-databases, they found that 91%, 86.5% and 79.2% of *Jatropha* ESTs were homologous to castor bean, poplar and *Arabidopsis*, respectively. The first-generation linkage map could lay a solid base for QTL mapping of agronomic traits, marker-assisted breeding and genetic improvement for improvement of the crop productivity. Roorkiwal *et al.* (2014) selected 10 abiotic stress responsive genes for allele specific sequencing across a chickpea reference set of 300 genotypes. A total of 1.3 Mbp sequence data was generated. The maximum number of SNPs (34) was observed in abscisic acid stress and ripening (ASR) gene including 22 transitions, 11 transversions and one tri-allelic SNP. These genes can be used for association analysis and enhancing abiotic stress tolerance through molecular breeding. Singh *et al.* (2019) reported world's first model web server for crop variety identification using >350 Indian wheat varieties and Axiom 35 K SNP chip data. Standard filtering and linkage disequilibrium approach were used to develop varietal signature in Linux using HTML, Java, PHP and MySQL. VIsTa (Variety Identification System of *Triticum aestivum*) (http://webtom.cabgrid. res.in/vista), is a server based tool for variety identification and can be used in distinctness, uniformity and stability testing of varieties, dispute resolution of sovereignty, and management of germplasm in the endeavour of wheat productivity.

Diversity arrays technology: Diversity arrays technology is an inexpensive, highly reproducible microarray hybridization technology. This type of technology removes some of the major hurdles associated with RFLP, AFLP and SSR. DArT is chiefly appropriate for genotyping large genome size polyploid species, such as wheat (Zhang *et al.* 2011). Akbari *et al.* (2006) showed that DArT markers like SSR, AFLP and RFLP generated similar information with same performance among bread wheat cultivars. The first genetic map of pigeonpea based on diversity array technology markers was developed by Yang *et al.* (2011). The total number of DArT marker loci segregated in a Mendelian manner was 405 with 73% of DArT markers having unique segregation patterns. Two groups of genetic linkage maps were produced with these DArT markers. The maternal genetic linkage map had 122 unique DArT maternal marker loci with length of 270 cM, whereas the paternal genetic linkage map has a total of 172 unique DArT paternal marker loci with length of 451.6 cM. The molecular linkage maps constructed by DArT markers will be useful in locating markers linked to genes of economic interest. Gemenet *et al.* (2015) presented the first report on genetic polymorphisms of phosphorus (P) related traits, flowering time (FLO) and grain yield (GY) characters under P-limiting conditions in 151 West African pearl millet inbred lines using 285 DArT markers. They reported 9 markers significantly associated with P-related traits, 9 markers with FLO, and 13 markers with GY. Such findings can fill the gap of studying complex traits between quantitative and molecular methods.

Next generation sequencing: The advancement through next generation sequencing technique will increase our knowledge about how the genotypic variation can be translated into phenotypic one. The list of NGS technologies presently available are Roche/454 FLX system, Illumina/ Solexa Genome Analyzer, Applied Biosystems SOLiD System, Helicos single-molecule sequencing, Pacific Biosciences SMRT instruments and Oxford Nanopore Technology. These technologies perform easily population-genetic studies based on complete genome sequences rather than just a short sequence of a single gene. DNA sequencing from next generation systems produce shorter read lengths, different error profiles and higher coverage than Sanger sequencing data. The NGS discover novel genetic variation in large number of genes and gene families from related species, allelic variation in close or distant relatives can be easily characterized, and it also supports rapid domestication of new plant species (Henry 2011).

Transcriptome assemblies of chickpea and pigeonpea was developed by using next-generation sequencing (NGS) and high-throughput (HTP) genotyping methods (Varshney *et al.* 2012). Based on Sanger, 454/FLX and Illumina transcript reads, transcriptome assemblies of chickpea contained 44,845 transcript assembly contigs (TACs) and pigeonpea contained 21,434 TACs. Illumina sequencing of parental genotypes has resulted in the development of 120 million reads for chickpea and 128.9 million reads for pigeonpea. Alignment of these Illumina reads with their respective transcriptome assemblies have provided >10,000 SNPs each in chickpea and pigeonpea. Later on Kudapa *et al.* (2014) developed a complete transcriptome assembly of chickpea containing 46,369 TACs using 134.95 million Illumina single-end reads, 7.12 million single-end FLX/454 reads and 139,214 Sanger expressed sequence tags (ESTs). The transcriptome assembly recognized novel markers that can be used for gene discovery, marker-trait association and diversity analysis to improve research and breeding applications in chickpea and other related legumes. The analysis of phenotyping data together with genotyping data has provided candidate markers for drought-tolerance in chickpea and fertility restoration in pigeonpea. Yadav *et al.* (2017) conducted an NGS study in *Populus deltoides* to see the actual response of genes to increased CO_2 level. Higher number of transcripts were expressed in treatment as compared to control. Differential gene expression analysis showed 1951 transcripts were down regulated, whereas 2603 transcripts were up regulated and 159,982 transcripts showed no significance. The study showed that plants growing in higher level of atmospheric CO_2 can change their adaptation, productivity and vegetation by changing number of genes and expression patterns. The information

obtained from NGS study will certainly accelerate crop improvement and productivity in such legumes.

Conclusion

With the changing climate, huge loss of the important plant species has triggered the conservation of plant resources from last few years. Thus to conserve plant resource, there is a need for accurately identifying and characterizing the plant materials to ensure their sustainable use. The use of molecular markers has revolutionized the techniques for characterizing genetic variation and validates genetic selection. However, molecular markers are the most reliable source for the analysis of genome structure and behaviour in crop plants. Molecular markers such as RAPD, AFLP, microsatellites and SNP are used for genetic analysis of crop plant. Further, organelle based markers of chloroplast, mitochondria and ribosomes serve well in phylogenetic analysis and DNA barcoding of important agricultural and horticultural crops. Each of the marker system has its own advantages and disadvantages. The limitation of one system can lead to the development of improved and new marker system with better ability for conserving genetic resources of a crop.

ACKNOWLEDGMENTS

We thank Department of Science and Technology-Science and Engineering Research Board (DST-SERB) (sanction order No. PDF/2017/001488), India for the financial support.

REFERENCES

Ahmed E A, Mohamed S, Abdlulaziz A A, Ahmed M K, Hamed A F and Najeb B M. 2013. Identification of new SRAP markers linked to leaf chlorophyll content, flag leaf senescence and cell membrane stability traits in wheat under water-stressed condition. *Australian Journal of Crop Science* **7**: 887–93.

Akbari M, Wenz P, Caig V, Carling J and Xia L. 2006. Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. *Theoretical and Applied Genetics* **113**: 1409–20.

Altintas S, Toklu F, Kafkas S, Kilian B, Brandolini A and Ozkan H. 2008. Estimating genetic diversity in durum and bread wheat cultivars from Turkey using AFLP and SAMPL markers. *Plant Breeding* **127**: 9–14.

Arzani A and Ashraf M. 2017. Cultivated ancient wheats (*Triticum* spp.): A potential source of health beneficial food products. *Comprehensive Reviews in Food Science and Food Safety* **16**: 477–88.

Basha S D and Sujatha M. 2007. Inter and intra-population variability of *Jatropha curcas* (L.) characterized by RAPD and ISSR markers and development of population-specific SCAR markers. *Euphytica* **156**: 375–86.

Benor S. 2018. Molecular phylogeny of the genus Corchorus (Grewioideae, Malvaceaes.l.) based on nuclear rDNA ITS sequences. *The Crop Journal* **6**: 552–63.

Binyamin R, Khan M A, Khan N A and Khan A I. 2015. Application of SCAR markers linked with mungbean yellow mosaic virus disease-resistance gene in Pakistan mungbean germplasm. *Genetics and Molecular Research* **14**: 2825–30.

Chial H. 2008. DNA fingerprinting using amplified fragment length polymorphisms (AFLP): No genome sequence required. *Nature Education* **1**: 176.

Choudhary P, Khanna S M, Jain P K, Bharadwaj C, Kumar J, Lakhera P C and Srinivasan R. 2013. Molecular characterization of primary gene pool of chickpea based on ISSR markers. *Biochemical Genetics* **51**: 306–22.

Clarridge J E. 2004. Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. *Clinical Microbiology Reviews* **17**: 840–62.

Collard B C Y and Mackill D J. 2009. Start Codon Targeted (SCoT) Polymorphism: A simple, novel DNA marker technique for generating gene-targeted markers in plants. *Plant Molecular Biology Reports* **27**: 86.

Daniell H, Lin C, Yu M and Chang W. 2016. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biology* **17**:134 DOI 10.1186/s13059-016-1004-2.

Dar A A, Choudhury A R and Arumugam N. 2014. A study on seed protein profile of Indian cultivars of *Sesamum indicum* L. *International Journal of Current Biotechnology* **2**: 10–17.

Dar A A, Mudigunda S, Mittal P K and Arumugam N. 2017a. Comparative assessment of genetic diversity in *Sesamum indicum* L. using RAPD and SSR markers. *3 Biotech* **7**: 10.

Dar A A, Mahajan R, Lay P and Sharma S. 2017b. Genetic diversity and population structure of *Cucumis sativus* L. by using SSR markers. *3 Biotech* **7**: 307.

Dhanasekar P, Dhumal K N and Reddy K S. 2010. Identification of RAPD markers linked to plant type gene in pigeonpea. *Indian Journal of Biotechnology* **9**: 58–63.

Dong W, Liu J, Yu J, Wang L and Zhou S. 2012. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. *PLoS One* **7**: e35071.

Edwards D and Batley J. 2010. Plant genome sequencing: applications for plant improvement. *Plant Biotechnology Journal* **8**: 2–9.

El-Esawi M A, Al-Ghamdi A A, Ali H M, Alayafi A A, Witczak J and Ahmad M. 2018. Analysis of genetic variation and enhancement of salt tolerance in French pea (*Pisum Sativum* L.). *International Journal of Molecular Sciences* **19**: 2433.

Esposito M A, Martin E A, Cravero V P and Cointry E. 2007. Characterization of pea accessions by SRAP's markers. *Scientia Horticulturae* **113**: 329–35.

Gao X, Zhu Y P, Wu B C, Zhao Y M and Chen J Q. 2008. Phylogeny of Dioscorea sect. Stenophora based on chloroplast matK, rbcL and trnL-F sequences. *Journal of Systematics and Evolution* **46**: 315–21.

Garris A J, Tai T H, Coburn J, Kresovich S and McCouch S. 2005. Genetic structure and diversity in *Oryza sativa* L. *Genetics* **169**: 1631–8.

Gemenet D C, Leiser W L, Zangre R G, Angarawai I I and Sanogo M D. 2015. Association analysis of low-phosphorus tolerance in West African pearl millet using DArT markers. *Molecular Breeding* **35**: 171 DOI 10.1007/s11032-015-0361-y

Guenni K, Aouadi M, Chatti K and Salhi-Hannachi A. 2016. Analysis of genetic diversity of Tunisian pistachio (*Pistacia vera* L.) using sequence related amplified polymorphism (SRAP) markers. *Genetics and Molecular Research* **15**: gmr15048760.

Gupta N, Zargar S M, Mahajan R, Sharma S, Nehvi F A and Salgotra R K. 2018. Candidate gene-based characterization of common bean genotypes. *Indian Journal of Genetics* **78**: 389–92.

Henry R J. 2011. Next-generation sequencing for understanding

and accelerating crop domestication. *Briefings in Functional Genomics* **2**: 51–6.

Hollingsworth P M, Graham S W, Damon P and Little D P. 2011. Choosing and using a plant DNA barcode. *PLoS One* **6**: e19254.

Kalia R K, Manoj K R, Sanjay K, Rohtas S and Dhawan A K. 2011. Microsatellite markers: An overview of the recent progress in plants. *Euphytica* **177**: 309–34.

Kress W J, Wurdack K J, Zimmer E A, Weigt L A and Janzen D H. 2005. Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences USA* **102**: 8369–74.

Kudapa H, Azam S, Sharpe A G, Taran B and Li R. 2014. Comprehensive transcriptome assembly of chickpea (*Cicer arietinum* L.) using Sanger and next generation sequencing platforms: Development and Applications. *PLoS ONE* **9**: e86039.

Kumar P, Gupta V K, Misra A K, Modi D R and Pandey B K. 2009. Potential of molecular markers in plant biotechnology. *Plant Omics Journal* **2**: 141–62.

Lee S J, Shin Y, Kim Y and Lee S. 2017. Molecular markers based on chloroplast and nuclear ribosomal DNA regions which distinguish Korean-specific ecotypes of the medicinal plant *Cudrania tricuspidata* Bureau. *Journal of Plant Biotechnology* 44: 235–42.

Lu G, Cao J, Yu X, Xiang X and Chen H. 2008. Mapping QTLs for root morphological traits in *Brassica rapa* L. based on AFLP and RAPD markers. *Journal of Applied Genetics* **49**: 23–31.

Luo C, He X, Chen H, Ou S, Gao M, Brown J S, Tondo C T and Schnell R J. 2011. Genetic diversity of mango cultivars estimated using SCoT and ISSR markers. *Biochemical Systematics and Ecology* **39**: 676–84.

Motawea M H, Said A A and Khaled A G A. 2015. ISSR markers-trait associations and stability analysis in bread wheat varieties. *Plant Breeding and Biotechnology* **3**:167–77.

Nirmala S, Mukesh Y, Venkataraman B K, Kumar S R and Kumar J P. 2016. Hybridization between salt resistant and salt susceptible genotypes of mungbean (*Vigna radiata* L. Wilczek) and purity testing of the hybrids using SSRs markers. *Journal of Integrative Agriculture* **15**: 521–7.

Obert D E, Skinner D Z and Stuteville D L. 2000. Association of AFLP markers with downy mildew resistance in autotetraploid alfalfa. *Molecular Breeding* **6**: 287–94.

Pandey P K, Dhotre D P, Dharne M S, Khadse A N and Hiremath U I. 2007. Evaluation of mitochondrial 12S rRNA gene in the identification of *Panthera pardus* fusca (Meyer, 1794) from field-collected scat samples in the Western Ghats, Maharashtra, India. *Current Science* **92**: 1129–33.

Paull J G, Chalmers K J, Karakousis A, Kretschmer J M, Manning S and Langridge P. 1998. Genetic diversity in Australian wheat varieties and breeding material based on RFLP data. *Theoretical and Applied Genetics* **96**: 435–46.

Que Y, Pan Y, Lu Y, Yang C, Yang Y, Huang N and Xu L. 2014. Genetic analysis of diversity within a Chinese local sugarcane germplasm based on start codon targeted polymorphism. *BioMed Research International* **2014**: 468375.

Robarts D W H and Wolfe A D. 2014. Sequence-related amplified polymorphism (SRAP) markers: A potential resource for studies in plant molecular biology. *Applications in Plant Science* 2: 1400017.

Roorkiwal M, Nayak S N, Thudi M, Upadhyaya H D and Brunel D. 2014. Allele diversity for abiotic stress responsive candidate genes in chickpea reference set using gene based SNP markers. *Frontiers in Plant Science* **5**: 248.

Schoch C L, Seifert K A, Huhndorf S, Robert V and Spouge J L. 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for *Fungi. Proceedings of National Academy of Sciences* **109**: 6241–6.

Sharma R, Dar A A, Mahajan R and Sharma S. 2019. Molecular and biochemical characterisation of Indian germplasm of *Pisum sativum* L. *Proceedings of the National Academy of Sciences of India Section B: Biological Sciences*.https://doi.org/10.1007/s40011-018-01069-3.

Sharma S, Pradhan S, Khajuria P and Dar A A. 2017. Genetic diversity studies in cucumber (*Cucumis sativus* L.) using RAPD markers. *Vegetable Science* **44**: 34–7.

Shu J, Liu Y, Li Z, Zhang L and Fang Z. 2016. Detection of the diversity of cytoplasmic male sterility sources in broccoli (*Brassica Oleracea* var. *Italica*) using mitochondrial markers. *Frontiers in Plant Science* 7: 927.

Singh R, Iquebal M A, Mishra C N, Jaiswal S and Kumar D. 2019. Development of model web-server for crop variety identification using throughput SNP genotyping data. *Scientific Reports* **9**: 5122.

Smith R L, Schweder M E, Chowdhury M K U, Seib J C and Hank S C. 1993. Development and application of RFLP and RAPD DNA markers in genetic improvement of pennzsetum for biomass and forage production. *Biomass and Bioenergy* **5**: 51–62.

Smykal P, Aubert G, Burstin J, Coyne C J and Ellis N T. 2012. Pea (*Pisum sativum* L.) in the genomic era. *Agronomy* **2**: 74–115.

Su M, Hsu T, Wang C, Lin K and Chiang M. 2017. Genetic diversity of a novel oil crop, *Camellia brevistyla* , revealed by ISSR DNA markers. *Horticultural Science and Technology* **35**: 588–98.

Suprasanna P, Sidha M and Ganapathi T R. 2008. Characterization of radiation induced and tissue culture derived dwarf types in banana by using a SCAR marker. *Australian Journal of Crop Science* **1**: 47–52.

Thakur B, Sharma S, Sharma I, Sharma P, Deepika and Zargar S M. 2018. Diversity analysis of pea genotypes using RAPD markers. *Legume Research* **41**: 196–201.

Thatikunta R, Sankar A S, Sreelakshmi J, Palle G and Leela C. 2016. Utilization of *in silico* EST–SSR markers for diversity studies in castor (*Ricinus communis* L.). *Physiology and Molecular Biology of Plants* **22**: 535–45.

Varshney R K, Kudapa H, Roorkiwal M, Thudi M and Pandey M K. 2012. Advances in genetics and molecular breeding of three legume crops of semi-arid tropics using next-generation sequencing and high-throughput genotyping technologies. *Journal of Bioscience* **37**: 811–20.

Wang CM, Liu P, Yi C, Gu K and Sun F. 2011. A first generation microsatellite- and SNP-based linkage map of Jatropha. *PLoS ONE* **6**: e23632.

Wang M L, Barkley N A and Jenkins T M. 2009. Microsatellite markers in plants and insects. Part I. Applications of biotechnology. *Genes Genomes Genomics* **3**: 54–67.

Wünsch A and Hormaza J I. 2002. Molecular characterisation of sweet cherry (*Prunus avium* L.) genotypes using peach [*Prunus persica* (L.) Batsch] SSR sequences. *Heredity* **89**: 56–63.

Yadav SK, Barthwal S and Ginwal HS. 2017. Response of *Populus deltoides* Bartr. ex Marsh genes under elevated CO2 through Next Generation Sequencing (NGS) [Abstract]. In: Abstracts of the NGBT conference held at Bhubaneswar, Odisha October 02–04. *Canadian Journal of Biotechnology* **1**: 196.

Yang S Y, Saxena R K, Kulwal P L, Ash G J and Dubey A. 2011.

The first genetic map of pigeon pea based on diversity arrays technology (DArT) markers. *Journal of Genetics* **90**: 103–09.

Yang Y, Yang T and Jing Z. 2015. Genetic diversity and taxonomic studies of date plum (*Diospyros lotus* L.) using morphological traits and SCoT markers. *Biochemical Systematics and Ecology* **61**: 253–9.

Yu X, Ren S, Zhao L, Guo J and Bao Y. 2018. Molecular mapping of a novel wheat powdery mildew resistance gene *Ml92145E8-9* and its application in wheat breeding by marker-assisted selection. *The Crop Journal* 6: 621–7.

Zhang L, Liu D, Guo, Yang W and Sun J. 2011. Investigation of genetic diversity and population structure of common wheat cultivars in northern China using DArT markers. *BMC Genetics* **12**: 42.