Molecular Characterization of Plant Genetic Resources

Mukesh Kumar Rana
Division of Genomic Resources, ICAR- NBPGR, New Delhi

Plant Genetic Resources (PGR) are one of the most valuable natural to develop new cultivars either with high yield, or better quality, or more adapted to abiotic stress, or more resistant to pest and pathogens. Germplasm conservation activities comprise maintenance, characterization, and evaluation of the genetic diversity within a species. Characterization is an important step as it determines the genetic identity of each accession in the germplasm collection conserved in the gene banks. Proper utilization of PGR depends much on this step of correct characterization. The main purpose of characterization is to establish the identity of accessions and to discern genetic relationship among genotypes. Characterization, at present is carried out either based on morphological traits or on molecular markers (biochemical and DNA markers). Morphology-based characterization has some limitations in the accurate identification of the accessions, such as limited number of traits to characterize (Rao 2004), highly heritable traits show no variation over much of the material studied (Karp et al. 1997), and trait expression is subjected to strong environmental influence, mainly in quantitative traits (Karp et al. 1997; Rao 2004). Biochemical markers such as isozymes, allozymes and storage proteins are effective in avoiding environmental influence, but they are unable to detect low levels of variation (Rao 2004) because they screen a small section of the genome. DNA-based techniques overcome both disadvantages as they identify polymorphism at DNA sequence level, therefore they are independent of environmental influence, and furthermore they sample the whole genome (Ovesná et al. 2002). DNA markers have been classified as dominant and co-dominant markers. Advantage of co-dominant markers over dominant markers is the differentiation between homozygous and heterozygous individuals. Data analysis and interpretation for both types of markers in PGR characterization will be different.

DNA markers in PGR characterization

DNA markers are indispensable tools for measuring the diversity of plant species. Low assay cost, affordable hardware, throughput, convenience and ease of assay development and automation are important factors when choosing a technology (Rafalski & Tingey, 1993). There are many marker techniques available for PGR characterization including RAPD, AFLP, SSR, SNP etc. Databases based on a large number of potential characters are readily available for inferring relationships using sequence data. The advantages of RAPD include its simplicity, low cost, rapid, use of arbitrary primers, no need of initial genetic or genomic information, and the requirement of only tiny quantities of target DNA. Disadvantages of this technique are dominant type and the lack of a prior knowledge on the identity of the amplification products which in turn creates problems with reproducibility and co-migration (Munthali et al. 1992; Lowe et al. 1996). The major advantage of the AFLP technique is the large number of polymorphisms that the method generates compared with other markers. The ability of AFLP to differentiate individuals in a population makes the technique useful for paternity analyses, gene-flow experiments and also for plant variety registration. However, the methodology of AFLP experiment and post-run data analysis are complex and time consuming compared with other markers like RAPD. The great advantage of microsatellite or Simple Sequence Repeat (SSR) analysis is the large number of polymorphisms that the method reveals. The ability of the method to differentiate individuals when a combination of loci is examined makes the technique very useful for gene-flow experiments, cultivar identification and paternity analyses (Hokanson et al. 1998). Major problem with the microsatellite relates with the initial screening of an organism for microsatellite library creation. The distribution and frequencies of SNPs are the key factors to understand molecular diversity between closely related populations and species (Tautz et al. 2010). However, this technique

has bi-allelic nature and less resolution power compared with multi-allelic microsatellites (Engle et al. 2006); though this shortcoming is overcome by its inherent capacity of scanning a large number of loci. SNP markers are best for characterizing and conserving the gene bank materials and the AFLP and the microsatellite markers are more suitable for diversity analysis and fingerprinting (Varshney et al. 2007). An ideal marker does not exist for use in all studies; rather a technique or techniques will be suited to a range of investigations. AFLP and microsatellites should not be considered appropriate for phylogenetic analyses above the species level. These markers are undoubtedly valuable tools for addressing population genetics and plant breeding issues, but for phylogeny reconstruction and taxonomy they could be problematic and sometimes even misleading, so they must be used with caution. Molecular genetics is a fast-moving field and new techniques are likely to be developed in the near future which will have their own strengths and limitations. Development of molecular technique based on error free database is another essential demand for easy assignment of unknown plant samples into appropriate taxa.

Molecular data analysis

Molecular markers have been extensively used in characterization of PGR. Many different ways for analyzing the data from these characterizations are available, depending on the specific interest of the study. All the analyses start with a binary matrix of presence or absence of bands. Two main kinds of analysis can be achieved: analysis of relationship among accessions in a group, and identification of accessions. Reliability of identification of accessions may be achieved by maximal probability of identical match by chance, and the efficiency may be quantified recording the number of resolved genotypes. The analysis of relationship can be performed within groups and between sub-groups of the whole group. Criteria for shaping sub-groups can be arbitrary or natural. A general overview of the genetic diversity within a group can be achieved by calculating some parameters depending on the type of marker (dominant or co-dominant). Relationship between pairs of accessions can be quantified by similarity or dissimilarity indexes. Selection of the most adequate index will depend on the type of markers, relatedness of the accession, and mating system of the species. Visualization of the relationship among accessions can be carried out by ordination and classification techniques

Molecular markers in gene bank management

Management of large germplasm collections in the gene banks is often costly, time-consuming, and labor-intensive. Therefore, it is urgent to build a core collection, which, as the representative germplasm resource of the entire germplasm collection, preserves the maximum genetic diversity and minimum repetition of a crop species (Frankel & Brown, 1984). Therefore, a core collection can improve germplasm selection and evaluation for curators and breeders, while maintaining a core set that representative of the genetic diversity of the entire germplasm collection. To date, dozens of core collections have been successfully established with the aid of molecular markers including *Solanum pimpinellifolim* (Rao et al. 2012), *Cucumis sativus* (Lv et al. 2012), *Cucumis melo* (Hu et al. 2014), *Malus* × *domestica* (Richards et al. 2009; Liang et al. 2015), and *Ziziphus jujuba* (Xu et al. 2016).

One of the possibilities to improve the composition of a collection is to identify and eliminate redundancies. The redundant genotypes, heterogeneous structure, and unavailable information on trait diversity affect the successful utilization of the genetic potential of these collections (Xu et al. 2016). Redundant accessions increase the cost of maintenance and may not be able to capture other diversity. Many genebanks around the globe are using molecular marker technology for identifying duplicates (van Treuren & van Hintum 2001, Maras et al. 2017). Information at genetic level is also very useful for designing sampling strategies. Genetic diversity if one could relate to geography can be used to

determine the collection sites, sample sizes as well as for in-situ or on-farm conservation. Diversity at intra or inter landraces is useful for sampling strategies. After conservation over the years, it is important to study the genetic integrity and needs regular monitoring. DNA markers offer a great advantage for such kind of study in the genebanks.

References

Brown, AHD (1989) Core collections: a practical approach to genetic resources management. Genome 31: 818-824.

Engle LJ, Simpson CL, Landers JE (2006) Using high-throughput SNP technologies to study cancer. Oncogene, 25:1594–1601.

Frankel, OH and Brown, AHD (1984) Plant genetic resources today: a critical appraisal (eds Holden, JHW, Williams, JT) 249–257 (George Allen & Unwin).

Hokanson SC, Szewc-McFadden AK, Lamboy WF, McFerson JR (1998) Microsatellite (SSR) markers reveal genetic identities, genetic diversity and relationships in a *Malus x domestica* Borkh. core subset collection. Theor Appl Genet 97:671–683.

Hu J et al. (2014) Microsatellite diversity, population structure, and core collection formation in melon germplasm. Plant Mol Biol Rep 33: 439-447.

Karp A, Kresovich S, Bhat K, Ayad W, Hodgkin T (1997) Molecular tools in plant genetic resources conservation: a guide to the technologies. IPGRI Technical Bulletin No. 2. International Plant Genetic Resources Institute, Rome, Italy.

Krauss SL (1999) Complete exclusion of nonsires in an analysis of paternity in a natural plant population using amplified fragment length polymorphism (AFLP) Mol Ecol: 8:217–226.

Liang W et al (2015) Genetic diversity, population structure and construction of a core collection of apple cultivars from Italian germplasm. Plant Mol Biol Rep 33: 458-473.

Lowe AJ, Hanotte O, Guarino L (1996) Standardization of molecular genetic techniques for the characterization of germplasm collections: The case of random amplified polymorphic DNA (RAPD) Plant Genet Resour Newsl, 107:50–54.

Lv J. et al. (2012) Genetic diversity and population structure of cucumber (*Cucumis sativus* L.). PLoS ONE 7, e46919.

Maras M , Sedlar A, Reid A, Božović V, Jovović Z, Meglič V and Dolničar P (2017) Genetic diversity and redundancy among potato accessions in the Montenegrin collection as revealed by microsatellite markers. American Journal of Potato Research 94: 306–313.

Munthali M, Ford-Lloyd BV, Newbury HJ (1992) The random amplification of polymorphic DNA for fingerprinting plants. PCR Methods Appl 1: 274–276.

Ovesná J, Poláková K, Leisová L (2002) DNA analyses and their applications in plant breeding. Czech J Genet Plant Breed 38: 29–40.

Rafalski JA and Tingey SV (1993) Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. Trends Genet 9:275–280.

Rao ES, Kadirvel P, Symonds RC, Geethanjali S and Ebert AW (2012) Using SSR markers to map genetic diversity and population structure of *Solanum pimpinellifolium* for development of a core collection. Plant Genet Resour 10: 38-48.

Rao N (2004) Plant genetic resources: advancing conservation and use through biotechnology. Afr J Biotechnol 3:136–145.

Richards, CM et al. (2009) Selection of stratified core sets representing wild apple (*Malus sieversii*). J Am Soc Hortic Sci 134: 228-235.

Tautz D, Ellegren H, Weigel D (2010) Next generation molecular ecology. Mol Ecol, 19:1–3.

van Treuren R and ThJL van Hintum (2001) Identification of intra-accession genetic diversity in selfing crops using AFLP markers: Implications for collection management. Genet Resour Crop Evol 48: 287-295.

Varshney RK, Chabane K, Hendre PS, Aggarwal RK, Graner A (2007) Comparative assessment of EST-SSR, EST-SNP and AFLP markers for evaluation of genetic diversity and conservation of genetic resources using wild, cultivated and elite barleys. Plant Soc 173: 638–649.

Xu, CQ et al. (2016) Identifying the genetic diversity, genetic structure and a core collection of *Ziziphus jujuba* Mill. Var. jujuba accessions using microsatellite markers. Sci Rep 6, 31503; doi: 10.1038/srep31503.