Application of different molecular markers in biotechnology

Berna Kocaman¹, Sevim Toy², Sevgi Marakli¹,³,⁴*

¹Department of Biotechnology, Institution of Science, Amasya University, Amasya/Turkey
²Department of Biology, Institution of Science, Amasya University, Amasya/Turkey
³Department of Medical Services and Techniques, Sabuncuoglu Serefeddin Health Services Vocational School, Amasya University/Turkey
⁴Amasya University, Central Research Laboratory, Amasya/Turkey

Abstract

Several molecular markers have been developed to explore genetic diversity, resistance against biotic/abiotic stress, disease, biomarker and evolutionary relationships in different genomes. These markers could be classified as sequence, hybridisation, PCR and retrotransposon based techniques. In this study, procedure and applications of AFLP (Amplified Fragment Length Polymorphism), ITS (Internal Transcribed Spacer), IRAP (Inter Retrotransposon Amplified Polymorphism), SSR (Simple Sequence Repeats), VNTR (Variable Number Tandem Repeats), SNP (single nucleotide polymorphism), CAPS (Cleaved Amplified Polymorphic Sequences), SCoT (Start Codon Targeted Polymorphism), SSCP (Single Strand Conformational Polymorphism) markers in plant, animal and human genomes were discussed.

1. Introduction

Markers are properties that could be used to distinguish intra- and/or inter-populations. When compare to morphological markers, molecular markers are more useful, more robust and independent of environmental conditions. Moreover, the strength of these markers has been increased with advances in next-generation sequencing technologies (Grover and Sharma, 2016; Nadeem et al., 2018). Many studies are using these markers in different genomes. However, plant genomes are mainly investigated to analyse genetic diversity by using these markers. Here, we present a detailed procedure and application of various molecular marker techniques in animal and human in addition to plant genomes.

¹Correspondence: sevgi.marakli@amasya.edu.tr
2. AFLP (Amplified Fragment Length Polymorphism)

AFLP is one of the PCR-based molecular marker technologies using DNA fragments after digestion with restriction enzymes and compares fingerprints to analyse differences in DNA sequences (Figure 1). The advantage of this method is as follows: no prior information about target genome, high reproducibility and sensitivity. AFLP technique has been used for genetic diversity within and between species, genetic map and evolutionary relationships. Moreover, transcriptomic variations and epigenetic studies such as DNA methylation are also analysed by using this technique (Paun and Schönswetter, 2012).

**Figure 1.** Procedure of AFLP molecular marker. 1. Genomic DNA is digested with combination of EcoRI (blue) and MseI (red) restriction enzymes. 2. EcoRI- (blue) and MseI (red)- specific adapters are ligated to the fragment ends. 3. In pre-amplification step; primers- specific to adapter sequences are used to amplification of fragments. 4. In selective PCR amplification step, selective nucleotides are added to EcoRI and MseI primers. 5. Amplification products are analysed by using denaturing polyacrylamide gels (Vuylsteke et al., 2007)

There are numerous reports to identify polymorphisms among plants via AFLP molecular marker. One of them was carried out by Moya-Hernández et al. (2018). They performed comparative analyses related to genetic diversity of *C. ficifolia* found in some regions of Mexico, reporting a total of 195 bands with 24.6% polymorphism. Similar to this study,
Ovesna et al. (2018) used AFLP markers to investigate diversity in a different plant, *S. sonchifolius*. They found higher polymorphism ratio (97.3%) among plants and reported that this ratio much higher than previous studies.

In some studies, molecular marker combinations have been used to better understand efficiency of markers to investigate polymorphism and comparison the results. Cao et al. (2019) performed one of these studies by using Simple Sequence Repeat (SSR) and (AFLP) markers. They investigated genetic diversity of *Pyrus pyrifolia* var. Nakai. As a result of SSR, AFLP and SSR+AFLP analyses, they observed rich genetic diversity in Nakai varieties. Similarly, Hadipour et al. (2020) also analysed AFLP and ISSR primer combinations in 67 *P. bracteatum* genotypes. Among populations, 52% for ISSR and 48% for AFLP polymorphism ratios were detected.

3. ITS (Internal Transcribed Spacer)

Ribosomal RNA gene and spacer regions have provided phylogenetic knowledge in prokaryotes and eukaryotes. Ribosomal DNA consists of coding regions (18S, 5.8S and 28S) together with two ITS regions and one NTS (non-transcribed spacer) (Wei et al., 2006) (Figure 2). ITS is a degraded region during maturation in ribosomal transcript.

**Figure 2.** ITS regions together with ribosomal subunits (Choudhary et al., 2015)

ITS due to more variation properties than ribosomal sequences have been commonly used for molecular markers to figure out phylogenetic analysis and/or identification of bacteria, plant and animal strains and/or species (Lee et al., 2017; Rocha et al., 2019; Ali et al., 2019;
Fazeli-Nasab et al., 2020). Pourahmad et al. (2019) performed sequence analysis of the mycobacterial 16S-23S ITS region to determine aquatic mycobacteria species, reporting effectiveness of this marker. In another study, Yu et al. (2020) evaluated the species specificity of ITS. For this purpose, they improved a determination method based on 500 ITS sequences. They concluded that the accuracy ratio of this procedure was 99.3% and 100% for species level and genus level, respectively in Streptococcus. A similar study was performed in a fungal community. Deng et al. (2020) investigated the specificity of ITS sequences by using *in silico* and even experimental analyses. They suggested that primer selection could affect the finding of amplicon-based metabarcoding studies due to different taxonomic levels.

4. IRAP (Inter Retrotransposon Amplified Polymorphism)

IRAP is a retrotransposon-based molecular marker technique, using primers face outwards from LTR (Long Terminal Repeats) regions. Single primer or different primers depending on LTR sequences are used for amplifying genomic DNA regions between retrotransposons (Kalendar and Schulman, 2006) (Figure 3). In this technique, high polymorphism levels could detect without DNA digestion, ligations or probe hybridisation.

![Figure 3. Procedure of IRAP molecular marker](http://www.biocenter.helsinki.fi/bi/genomedynamics/markers.html)

There are many retrotransposons based molecular markers but IRAP method has been used for several studies to investigate genetic diversity, especially plant species (Noormohammadi et al, 2018; Lanciková and Žiarovská, 2020). Furthermore, there are also different studies
using combination of two retrotransposon markers: IRAP and REMAP (Retrotransposon-Microsatellite Amplified Polymorphisms). One of them was carried out by Holasou et al. (2019) to evaluate genetic diversity in wheat (*Triticum aestivum* L.). They reported that both methods produced highly polymorphic bands in samples. Shingote et al. (2019) also identified that IRAP system was superior to ISSR in terms of marker in terms of index, resolving power and polymorphic loci per assay.

In addition to REMAP and ISSR, SCoT (start codon-targeted) markers together with IRAP has also been used to analyse relationships among samples. Guan et al. (2020) investigate genetic diversity of 268 *Diospyros* accessions from different regions in China, determining 90 and 97 polymorphic alleles from nine SCoT and nine IRAP markers. Shehata et al. (2020) performed different study in both yeast (*Saccharomyces cerevisiae* L.) and barley (*Hordeum vulgare* L.) by using IRAP and SCoT. Obtaining findings indicated that different band patterns observed between control and salt treatments, and even the high levels of salinity could cause new retrotranspositions.

5. SSR (Simple Sequence Repeats)

Genomes consist of repetitive elements classified as interspersed repeats and tandem repeats depending on distribution SSR (microsatellites) with units of 1 to 6–10 bp are subclass of microsatellites in tandem repeats (Dumbovic et al., 2017) (Figure 4). There are two main approaches to develop SRR primers: i. Analysing known SSR primers already improved for related species and ii. production of genomic library and improving SSRs b using NGS technologies (Csencsics et al., 2010). In addition to genomic analyses, de novo transcriptome sequencing (RNA-Seq) is also reliable approach for SSR development in different species, (Taheri et al., 2018).

![Image of SSR markers](https://www.ncbi.nlm.nih.gov/probe/docs/techsts)
SSR markers have also been widely used in genetic diversity studies especially in plants such as rice (Jasim Aljumaili et al., 2018), walnut (Bernard et al., 2018), torch ginger (Ismail et al., 2019), maize (Adu et al., 2019), Norway spruce (Bínová et al., 2020) and chickpea (Asadi et al., 2020). Moreover, development of EST-SSR markers have been also studied in Indian mulberry (Thumilan et al., 2016), Lycium barbarum (Chen et al., 2017), Bletilla striata (Xu et al., 2018), Lilium (Biswas et al., 2018), Chinese Hawthorn (Ma et al., 2019) and opium poppy (Vašek et al., 2020). In addition to plants, these markers are also used for animal (Li et al., 2020; Silva Junior et al., 2020) and human (Pai et al., 2016).

6. VNTR (Variable Number Tandem Repeats)

VNTR also known as minisatellites are a member of repetitive DNA sequences dispersed in genome. They organised as tandem repeat units of a 10–60 base motif, flanked by conserved DNA restriction sites (Figure 5). These sequences show variations in length (number of repeats) among individuals (Singh et al., 2008).

Figure 5. SSR markers (https://www.ncbi.nlm.nih.gov/probe/docs/techsts)

VNTRs are an important source of RFLP (Restriction Fragment Length Polymorphism) markers in linkage analysis (mapping) of genomes. Similar to other molecular markers, VNTR method has been also utilised for genetic diversity and understand evolutionary relationships between species (Apablaza et al., 2015; Hu et al., 2015; Ghielmetti et al., 2017). The versatility of this marker allows it to be easily used in different organisms. Mathema et al.
(2019) evaluated six microsatellite markers were developed and tested in 37 *P. malariae* isolates, concluding sufficient heterozygosity among samples. Furthermore, Sneideris et al. (2019) reported high genetic diversity as a result of VNTR tests on *F. graminearum* isolates. Najar-Peerayeh et al. (2019) studied with other basteria species, *A. baumannii*, determining high level of polymorphism similar to Sneideris et al. (2019).

7. SNP (Single Nucleotide Polymorphism)

Single nucleotide polymorphisms (SNP) are found almost all living things. If more than 1% of a population not carry a same nucleotide at a specific DNA position, then this variation can be named as a SNP. SNPs can occur once every 1000 bases or so. Not only a SNP is a nucleotide difference in DNA but also causes genetic polymorphisms among individuals and even populations (Perkel, 2008) (Figure 6).

![Figure 6. SNPs are DNA differences found in a specific location](https://genetics.thetech.org/ask-a-geneticist)

SNPs are used to identify gene or genes responsible for traits (Li et al., 2020; Ayala-Usma et al., 2020). In human, animal and plant genomes, there are many studies in terms of SNPs analyses. Among them, Kuhn et al. (2019) investigated biallelic SNP markers in mango germplasm. Zafar et al. (2020) studied with the relationships between oleic acid and SNPs in *Brassica napus*. Moreover, Amanullah et al. (2020) developed SNPs based cleaved amplified polymorphism sequence (CAPS) markers in melon. In a different analysis, Zhao et al. (2019) studied with meat from yak and cattle, identifying higher polymorphism in the cattle population and monomorphism in the yak population. Fatai et al. (2020) performed *in silico*
analyses to identify functional and structural effects of inhibin A gene which is a growth factor and relationships between this gene and SNPs. In addition to correlating trait and SNPs, some SNPs can also be associated with certain diseases (Liu et al., 2020; Shibeshi et al., 2020).

8. CAPS (Cleaved Amplified Polymorphic Sequences)

CAPS are DNA fragments amplified by PCR using specific primers, then digestion of these amplicons with a restriction enzyme. Length polymorphisms as a result of variation in restriction sites are determined by gel electrophoresis (Figure 7). Therefore, this technique is also referred to PCR-Restriction Fragment Length Polymorphism (PCR-RFLP) (Singh et al., 2008).

Figure 7. Procedures of CAPS technique (Agarwal et al., 2008)

Polymorphisms analyses by using CAPS markers can be used for analysing structural properties (Babu et al., 2017), biotic/abiotic stress (Hubhachen et al., 2020) in addition to genetic diversity and molecular breeding (Kawahara et al., 2020). In one study, Gunaydin and Kafkas (2017) characterised strawberry varieties by using SSR and CAPS markers. They suggested that SSR was considered to be better than CAPS in terms of polymorphism rates.
Amanullah et al. (2020) combined SNP-CAPS markers to investigate in *Cucumis melo* L, identifying 7 QTLs for melon ovary traits.

9. SCoT (Start Codon Targeted Polymorphism)

SCoT is depended on short conserved region in genes surrounding the ATG translation start (or initiation) codon. Therefore, primers are designed according to the short conserved region flanking the ATG start codon (Figure 8). Amplicons are evaluated by standard gel electrophoresis with agarose gels (Collard and Mackill, 2009).

![Figure 8. Principle of SCoT analysis (Collard and Mackill, 2009)](image)

Similar to other molecular markers, SCoT could be successfully practised for genetic diversity (Vivodík et al., 2019; Vanijajiva et al., 2020; Xiao et al., 2020). In some studies, SCoT markers were combined with other molecular markers. Etminan et al. (2016) investigated SCoT and ISSR markers together to analyse durum wheat genotypes. In another study, Shekhawat et al. (2018) used SCoT and CBDP (CAAT-box derived polymorphism) markers for *Prosopis cineraria*. Similarly, Gholamian et al. (2019) also studied with *T. urartu* by using SCoT and CBDP markers. El-Fiki and Adly combined RAPD and SCoT markers to characterise potato cultivars.
10. SSCP (Single Strand Conformational Polymorphism)

SSCP is based on the electrophoretic mobility of a single-stranded DNA molecule in a non-denaturing gel. This conformational difference depends on the length of strand, sequence, location and number of base pairs (Gasser et al., 2006) (Figure 9).

**Figure 9.** The principle of SSCP analysis. Dot indicates point mutation cause single-strand conformations (Gasser et al., 2006)

Youssef and Shalaby (2016) studied with *Citrus Tristeza Virus* (CTV), an important virus for citrus, to characterise populations of CTV and determine haplotypes in populations by using SSCP marker. Zheng et al. (2016) also screened CRISRP-Cas9-mediated targeted mutagenesis in rice via SSCP application. They determined small indels and multiple mutants. In another study, Tchouomene-Labou et al. (2020) used SSCP technique to explore genetic differentiation, gene flow, demographic history and phylogenetic relationship in mitochondrial genes of *Glossina palpalis palpalis* populations.
11. Conclusion

Each marker system has advantages and disadvantages. Therefore, many studies have been performed by using combination of different markers to eliminate disadvantages. Moreover, the continuous development of molecular markers with advancements in sequencing technologies has provided a powerful tool in biotechnological research. It is important to know procedures and applications of molecular markers to improve existing markers and production of new markers.

References


Apablaza, P., Brevik, Ø. J., Mjøs, S., Valdebenito, S., Ilardi, P. et al. 2015. Variable Number of Tandem Repeats (VNTR) analysis of Flavobacterium psychrophilum from salmonids in Chile and Norway, BMC Veterinary Research, 11(1): 150.


Li, J., Hou, L., Sun, Y., Xing, J., Jiang, Y. et al. 2020. Single nucleotide polymorphism rs737028527 (G> A) affect miR-1b-3p and effects on chicken egg-laying traits, Animal Reproduction Science, 106476.


Liu, Z., Yin, X., Mai, H., Li, G., Lin, Z. et al. 2020. SCD rs41290540 single-nucleotide polymorphism modifies miR-498 binding and is associated with a decreased risk of coronary artery disease, Molecular Genetics & Genomic Medicine, 8(3): e1136.


Najar-Peerayeh, S., Karmostaji, A. 2019. Evaluation of multilocus Variable-Number Tandem-Repeat (MLVA-8 Orsay) for Typing of carbapenem-resistant Acinetobacter baumannii isolated from patients in Tehran, Iran, Archives of Clinical Infectious Diseases, 14(1): e64402s


Pai, T.W., Chen, C. M. 2016. SSRs as genetic markers in the human genome and their observable relationship to hereditary diseases, Biomarkers in Medicine, 10(6): 563-566.


Wei, N.W.V., Wallace, C.C., Dai, C.F., Pillay, K.M., Chen, C.A. 2006. Analyses of the ribosomal Internal Transcribed Spacers (ITS) and 5.8 S gene indicate that extremely high rDNA heterogeneity is a unique feature in the scleractinian coral genus Acropora (Scleractinia; Acroporidae), Zoological Studies, 45(3): 404.


