www.sciencevision.in

Colloquium

Sci Vis 11 (3), 116-124
July-September, 2011
ISSN (print) 0975-6175
ISSN (online) 2229-6026

Random amplified polymorphic DNA (RAPD) markers and its applications

N. Senthil Kumar[1] and G. Gurusubramanian[2]

[1]*Department of Biotechnology,* [2]*Department of Zoology, Mizoram University, Aizawl 796 004, India*

ABSTRACT

In the last decade, the random amplified polymorphic DNA (RAPD) technique based on the polymerase chain reaction (PCR) has been one of the most commonly used molecular techniques to develop DNA markers. RAPD is a modification of the PCR in which a single, short and arbitrary oligonucleotide primer, able to anneal and prime at multiple locations throughout the genome, can produce a spectrum of amplification products that are characteristics of the template DNA. RAPD markers have found a wide range of applications in gene mapping, population genetics, molecular evolutionary genetics, and plant and animal breeding. This is mainly due to the speed, cost and efficiency of the technique to generate large numbers of markers in a short period compared with previous methods. Therefore, RAPD technique can be performed in a moderate laboratory for most of its applications. It also has the advantage that no prior knowledge of the genome under research is necessary.

Key words: DNA; PCR; RAPD; polymorphism.

INTRODUCTION

Advances in molecular biology techniques have provided the basis for uncovering virtually unlimited numbers of DNA markers.[1] Over the last decade, polymerase chain reaction (PCR) has become a widespread technique for several novel genetic assays based on selective amplification of DNA.[2] This popularity of PCR is primarily due to its apparent simplicity and high probability of success. Unfortunately, because of the need for DNA sequence information, PCR assays are limited in their application. The discovery that PCR with random primers can be used to amplify a set of randomly distributed loci in any genome facilitated the development of genetic markers for a variety of purposes.[3,4]

RAPD (pronounced 'rapid'), for *R*andom *A*mplification of *P*olymorphic *D*NA, is a type of PCR reaction, but the segments of DNA that are amplified are random. The RAPD analysis described by Williams *et al*.[3] is a commonly used molecular marker in genetic diversity studies. No knowledge of the DNA sequence for the targeted gene is required, as the primers will bind somewhere in the se-

Proceedings of the "National Level Workshop on Random Amplified Polymorphic DNA (RAPD) Markers and It's Applications" organized on 20-21 May 2011 by the Departments of Biotechnology & Zoology, Bioinformatics Infrastructure Facility & State Biotech Hub, Mizoram University, in association with MIPOGRASS, and sponsored by DBT.
For correspondence: nskmzu@gmail.com (NSK)
gurus64@yahoo.com (GG)
Phones. +91-0389-2330859 (NSK)/2331021 (GG)

quence, but it is not certain exactly where. This makes the method popular for comparing the DNA of biological systems that have not had the attention of the scientific community, or in a system in which relatively few DNA sequences are compared (it is not suitable for forming a DNA databank). Perhaps the main reason for the success is the gain of a large number of genetic markers that require small amounts of DNA without the requirement for cloning, sequencing or any other form of the molecular characterisation of the genome of the species in question.

RAPD markers are decamer (10 nucleotide length) DNA fragments from PCR amplification of random segments of genomic DNA with single primer of arbitrary nucleotide sequence and which are able to differentiate between genetically distinct individuals, although not necessarily in a reproducible way. It is used to analyze the genetic diversity of an individual by using random primers. In this paper, the principles, working mechanism, differences between standard PCR and RAPD -PCR, characteristics, laboratory steps, data analysis and interpretation, advantages and disadvantages and several of the most common applications of RAPD markers in biology are discussed.

PRINCIPLE OF THE RAPD TECHNIQUE

The principle is that, a single, short oligonucleotide primer, which binds to many different loci, is used to amplify random sequences from a complex DNA template. This means that the amplified fragment generated by PCR depends on the length and size of both the primer and the target genome. The assumption is made that a given DNA sequence (complementary to that of the primer) will occur in the genome, on opposite DNA strands, in opposite orientation within a distance that is readily amplifiable by PCR. These amplified products (of up to 3.0 kb) are usually separated on agarose gels (1.5-2.0%) and visualised by ethidium bromide staining. The use of a single decamer oligonucleotide promotes the generation of several discrete DNA products and these are considered to originate from different genetic loci. Polymorphisms result from mutations or rearrangements either at or between the primer binding sites and are detected as the presence or absence of a particular RAPD band (Fig. 1). This means that RAPDs are dominant markers and, therefore, cannot be used to identify heterozygotes.

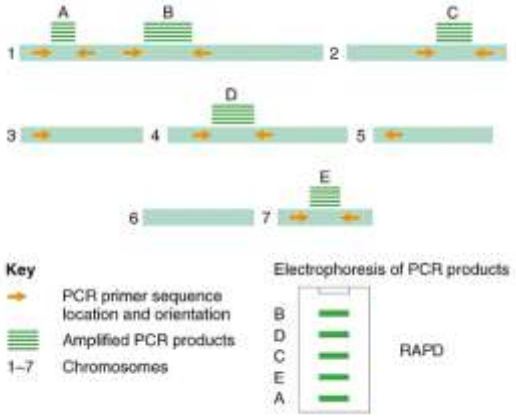


Figure 1. RAPD analysis general model [www.usask.ca/…/pawlin/ resources/rapds.html] (adapted from Griffiths *et al.*[6])

The standard RAPD utilises short synthetic oligonucleotides (10 bases long) of random sequences as primers to amplify nanogram amounts of total genomic DNA under low annealing temperatures by PCR. Primers are commercially available from various sources (e.g. Operon Technologies Inc., California; Biosciences, Bangalore; Eurofinns, Bangalore; GCC Biotech, Kolkata).

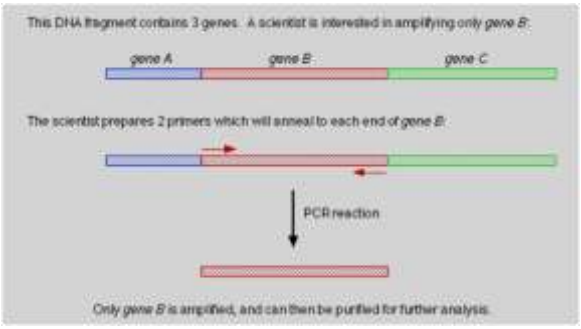
Welsh and McClelland[4] independently developed a similar methodology using primers about 15 nucleotides long and different amplification and electrophoretic conditions from RAPD and called it the arbitrarily primed polymerase chain reaction (AP-PCR) technique. PCR amplification with primers shorter than 10 nucleotides [DNA amplification finger-

printing (DAF)] has also been used to produce more complex DNA fingerprinting profiles.[5] Although these approaches are different with respect to the length of the random primers, amplification conditions and visualization methods, they all differ from the standard PCR condition[2] in that only a single oligonucleotide of random sequence is employed and no prior knowledge of the genome subjected to analysis is required.[6]

At an appropriate annealing temperature during the thermal cycle, oligonucleotide primers of random sequence bind several priming sites on the complementary sequences in the template genomic DNA and produce discrete DNA products if these priming sites are within an amplifiable distance of each other (Fig. 1). The profile of amplified DNA primarily depends on nucleotide sequence homology between the template DNA and oligonucleotide primer at the end of each amplified product. Nucleotide variation between different sets of template DNAs will result in the presence or absence of bands because of changes in the priming sites. Recently, sequence characterised amplified regions (SCARs) analysis of RAPD polymorphisms[7,8] showed that one cause of RAPD polymorphisms is chromosomal rearrangements such as insertions/deletions. Therefore, amplification products from the same alleles in a heterozygote differ in length and will be detected as presence and absence of bands in the RAPD profile, which is similar to that of low stringency minisatellite DNA fingerprinting patterns and is therefore also termed RAPD fingerprinting. On average, each primer directs amplification of several discrete loci in the genome so that allelism is not distinguishable in RAPD patterns.

DIFFERENCES BETWEEN STANDARD PCR AND RAPD PCR

Standard PCR



RAPD PCR

In RAPD analysis, the target sequence(s) (to be amplified) is unknown. A primer is designed with an arbitrary sequence. In order for PCR to occur: 1) the primers must anneal in a particular orientation (such that they point towards each other) and, 2) they must anneal within a reasonable distance of one another. Figure 2 depicts a RAPD reaction, a large fragment of DNA (genome A) is used as the template in a PCR reaction containing many copies of a single arbitrary primer.

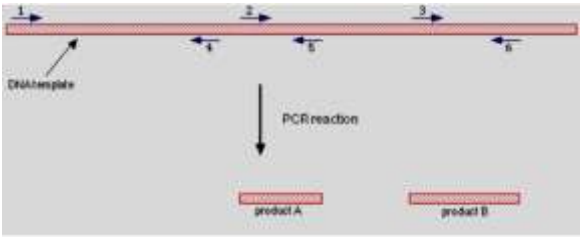


Figure 2. RAPD reaction for genome A. The arrows represent multiple copies of a primer (all primers have the same sequence). The direction of the arrow also indicates the direction in which DNA synthesis will occur. The numbers represent locations on the DNA template to which the primers anneal. Primers anneal to sites 1, 2, and 3 on the bottom strand of the DNA template and primers anneal to sites 4, 5, and 6 on the top strand of the DNA template.

RAPD reaction #1 for genome A. In the above example (Fig. 2), only 2 RAPD PCR products are formed:

1) Product A is produced by PCR amplification of the DNA sequence which lies in be-

tween the primers bound at positions 2 and 5.

2) Product B is produced by PCR amplification of the DNA sequence which lies in between the primers bound at positions 3 and 6.

Note that no PCR product is produced by the primers bound at positions 1 and 4 because these primers are too far apart to allow completion of the PCR reaction. Also no PCR products are produced at positions 4 and 2 or positions 5 and 3 because these primer pairs are not oriented towards each other.

FINDING DIFFERENCES BETWEEN GENOMES USING RAPD ANALYSIS

Consider the Figure 2 (genome A). If another DNA template (genome B) was obtained from a different (yet related) source, there would probably be some differences in the DNA sequence of the two templates. Suppose there was a change in sequence at primer annealing site #2 (Fig. 3):

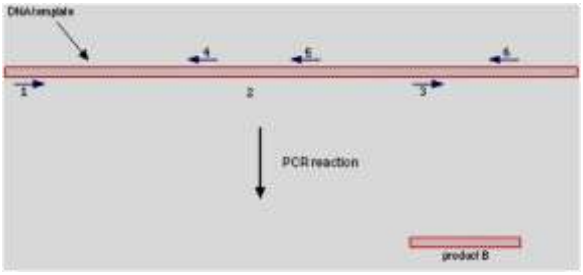


Figure 3. RAPD reaction for genome B. Primers anneal to sites 1, 2, and 3 on the top strand of the DNA template and primers anneal to sites 4, 5, and 6 on the bottom strand of the DNA template.

RAPD reaction #2 for genome B. As shown in Figure 3, the primer is no longer able to anneal to site #2, and thus the PCR product A is not produced. Only product B is produced. If you were to run the 2 RAPD PCR reactions diagramed above (genomes A and B) on an agarose gel, this is what you would see in Figure 4.

Genome A and B can represent genomic DNA from two individuals in the same species or possibly from two different species. Certain portions of genomic DNA tend to much conserved (very little variation) while other portions tend to vary greatly among individuals within a species or among different species. The trick in RAPD PCR analysis is to: 1) find those sequences which have just enough variation to allow us to detect differences among the organisms that we are studying and 2) find the right PCR primers which will allow us to detect sequence differences.

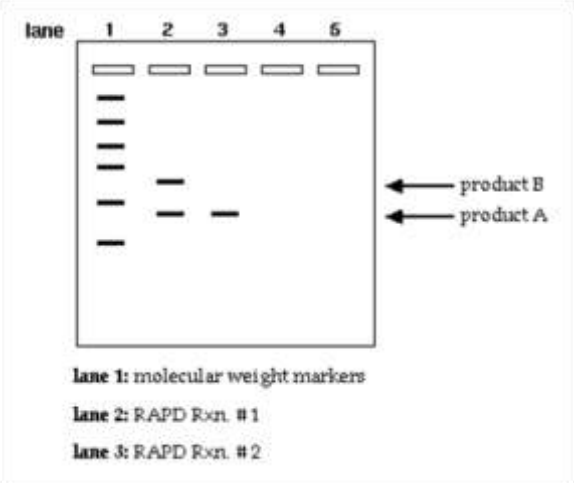


Figure 4: RAPD reactions on an agarose gel.

INTERPRETING RAPD BANDING PATTERNS

Each gel is analyzed by scoring the present (1) or absent (0) polymorphic bands in individual lanes. The scoring can be done based on the banding profiles which is clear and transparent (Fig. 5a) otherwise the scoring is very difficult (Fig. 5b).

Because of the nature of RAPD markers, only the presence or absence of a particular band can be assessed. Criteria for selecting scoring bands: 1) reproducibility—need to repeat experiments 2) thickness 3) size and 4) expected segregation observed in a mapping population. DNA polymorphism among individuals can be due to: 1) mismatches at the

primer site 2) appearance of a new primer site and 3) length of the amplified region between primer sites.

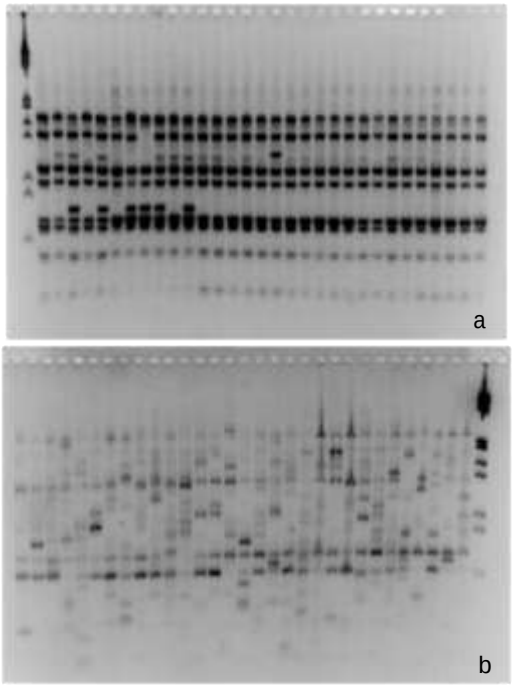


Figure 5. a. Very high quality RAPD gel. Both presence and absence of most bands are very clear and the background is transparent. Hence, scoring is easy; b. The bands are fuzzy. Those at the top have a smear starting from the well where the PCR product was loaded and many are observed only with difficulty. Only some bands are clear and can be scored (Courtesy: IPGRI and Cornell University, 2003).

The NTSYS-pc software ver. 2.02 is used to estimate genetic similarities with the Jaccard's coefficient.[9] The matrix of generated similarities is analyzed by the unweighted pairgroup method with arithmetic average (UPGMA), using the SAHN clustering module. The cophenetic module is applied to compute a cophenetic value matrix using the UPGMA matrix. The MXCOMP module is then used to compute the cophenetic correlation, i.e. to test the goodness of fit of the cluster analysis to the similarity matrix.

CHARACTERISTICS OF RAPD MARKER IN COMPARISON WITH OTHER NUCLEAR DNA MARKERS

Several types of molecular markers are available but none of them can be regarded as optimal for all applications.[10] The characteristic features of RAPD nuclear DNA marker[11] is summarized in Table 1.

Table 1. Features of RAPD nuclear DNA marker.

Features	RAPD
Allelic information*	Dominant
Locus presentation**	Multi–locus
DNA required (µg)	0.02
PCR–based	Yes
Restriction digestion	No
Reproducibility	Low
Development cost	Low
Cost per assay	Low
Suitability	Do not require prior molecular information
	Simple method and is the least expensive

* Dominant markers can identify only one allele (presence or absence of a band) and are therefore unable to determine heterozygosity; co-dominant markers are able to identify both the alleles.
** Multi-locus markers can visualize many genes simultaneously in contrast to only one region amplification by single-locus markers; however, the latter can easily be multiplexed for more reliable fingerprinting.

Laboratory steps for RAPD

Equipments required. Refrigerator and freezer, laminar flow hood, centrifuge, thermocycler, power supply units, hotplate or microwave, pH meter, standard balance, gel electrophoresis units, UV transilluminator.

Extraction of DNA. Tiny amounts of total, chloroplast or mitochondrial DNA can be

used. DNA must be clean and of high molecular weight. If minimal quality of DNA is not achieved, the reproducibility of results will be hard to ensure.[12]

Determination of DNA contents. Make a 10-100-fold dilution.[12] Measure the absorbance at 260 and 280 nm. The ratio between OD 260/OD 280 should not be less than 1.6. Calculation of DNA concentration. For example, OD 260 of a DNA sample is 0.020. OD 260 = 1 is equivalent to double strand DNA 50 µg/ml. OD 260 = 0.020 is equivalent to double strand DNA 50x0.020 = 1.0 µg/ml.

PCR reaction with a primer.[13,14] The PCR reaction mixture contains –

1. *Template DNA.* The quality of the template influences the outcome of the PCR. For instance, large amounts of RNA in a DNA template can chelate Mg^{2+} and reduce the yield of PCR. Also, impure templates may contain polymerase inhibitors that decrease the efficiency of the reaction. The integrity of the template is also important. Template DNA should be of high molecular weight. To check the size and quality of the DNA, run an aliquot on an agarose gel. When testing a new template, always include a positive control with primers that amplify a product of known size and produce a good yield.[13] Also include a negative control (without template DNA) to rule out any contaminants. The amount of template in a reaction strongly influences performance in PCR. The recommended amount of template for RAPD PCR is:

   Plant: 50 ng (but it may vary from species to species, so we have to optimize the template DNA concentration for each plant of our desire)

   Bacterial DNA: 1-10 ng

   Plasmid: 0.1-1 ng

   Human genomic DNA: 500 ng (maximum)

2. *Primers.* Use 10 base primers for PCR. Only 1 primer per reaction. Short primers bind randomly on the chromosomes. In most PCR amplifications, it is the sequence and the concentration of the primers that determine the overall assay success. A primer which brings about polymorphism between the samples to be tested considered good.
3. *Taq DNA polymerase.* It is obtained from hot spring bacterium, *Thermus aquaticus*. For most assays, the optimal amount of thermostable DNA polymerase (or a blend of polymerase) should be between 0.5-2.5 U/50 µl reaction volume. Increased enzyme concentration sometimes leads to decrease specificity.
4. *$MgCl_2$.* The optimal Mg^{2+} concentration may vary from approximately 1-5 mM. The most commonly used mg^{2+} concentration is 1.5 mM (with dNTPs at a concentration of 20 µM). Mg^{2+} influence enzyme activity and increases the temperature of double stranded DNA; excess Mg^{2+} in the reaction can increase non-specific primer binding and increase the non-specific background of the reaction.[14]
5. *dNTPs*. All four dNTPs should be balanced to minimize polymerase error rate. Imbalanced dNTPs mixtures will reduce Taq DNA polymerase fidelity. Increase in the concentration of dNTPs will require increase in Mg^{2+} concentration. Increase in dNTP concentration reduces free Mg^{2+}, thus interfering with polymerase activity and decreasing primer annealing. The final dNTPs concentration should be 50-500 µM (each dNTP) and the most commonly used is 200 µM .[14]
6. *PCR buffer (pH)*. Generally, the pH of the reaction buffer supplied with the corresponding thermostable DNA polymerase (pH 8.3-9.0) will give optimal results. However, for some systems, raising the pH may stabilize the template and enhance results.

DNA PCR Amplification.

1. Preparation of master mix for PCR. To a

PCR tube, add all the ingredients in order:

Solution	Stock concentration	Working concentration	Volume (µl)/ reaction	volume (µl) / 10 reactions
DNA		50 ng/µl	1	10
PCR buffer	10X	1X	1	10
MgCl2	50 mM	5 mM	1.5	15
dNTP mixture	200 mM	10 mM	0.75	7.5
Primer	500 pmole/µl	10 pmole/µl	0.3	3
Taq DNA polymerase	3 unit	3 unit	0.075	2.275
Autoclaved MQ			10.375	100.375
	Total		15	150

2. Tap the tube for 1-2 seconds to mix the contents thoroughly.
3. Add 25 µl of mineral oil in the tube to avoid evaporation of the contents.
4. Place the tube in the thermocycler block and set the program to get DNA amplification.
5. DNA amplification cycle: carry out the amplification in a thermocycler for 30-40 cycles using the following reaction conditions:

Initial denaturation at 94°C for 10 minutes
↓
Denaturation at 94°C for 1 min.
↓
Annealing at 37-45°C for 1 min.
↓
Extension at 72°C for 1 min.
↓
Final extension at 72°C for 10 minutes
↓
Cooling at 4°C.

Separating DNA fragments by Agarose gel electrophoresis.

1. 10X TBE buffer (108 g Tris base, 55 g Boric acid, 40 ml of 0.5 M EDTA (pH 8.0), make up to 1 L with water).
2. 1X TBE buffer (to prepare 500 ml of 1X TAE buffer, add 10 ml of 50X TAE buffer to 490 ml of sterile distilled water. Mix well before use).
3. DNA loading dye (bromophenol blue - 0.25%, xylene cyanol - 0.25%, Sucrose - 40% (w/v))

Weigh 25 mg of bromophenol blue, 25 mg of xylene cyanol and 4 g of sucrose and dissolve in 5 ml of distilled water. Make up to 10 ml by adding distilled water and sterilize. Store at -4°C.

Method:

- Prepare 50 ml of 0.8% agarose gel by adding 0.4 g agarose to 50 ml of 1X TAE buffer in a conical flask.
- Heat the mixture on a microwave, swirling the conical flask, until the agarose dissolves completely.
- Allow to cool down to 55-60°C.
- Add 0.5 µl ethidium bromide, mix well and pour the gel solution into the gel tray. Allow the gel to solidify for about 30 minutes at room temperature.
- Load 5 µl of ready to use DNA ladder into the first well. Add 2 µl of 6X gel loading buffer to 10 µl of PCR product. Load the PCR product into the following wells.
- Connect the power cord to the electrophoretic power supply according to the conventions: red-anode and black-cathode.
- Apply a voltage of 100-120 volts and 90 mA.
- Stop the electrophoresis, when bromophenol blue crosses more than 2/3 of the length of the gel.

Visualizing DNA fragments, using ethidium bromide. Visualize DNA bands on an UV trans-illuminator. If require, this profile can be photographed through a red filter for documentation using Gel Doc system.

ADVANTAGES OF RAPD

RAPD has been used widely because of the following advantages:

- It requires no DNA probes and sequence information for the design of specific primers.
- It involves no blotting or hybridisation steps, hence, it is quick, simple and efficient.
- It requires only small amounts of DNA (about 10 ng per reaction) and the procedure can be automated.
- High number of fragments.
- Arbitrary primers are easily purchased.
- Unit costs per assay are low compared to other marker technologies.

DISADVANTAGES OF RAPD

- Nearly all RAPD markers are dominant, i.e. it is not possible to distinguish whether a DNA segment is amplified from a locus that is heterozygous (1 copy) or homozygous (2 copies). Co-dominant RAPD markers, observed as different-sized DNA segments amplified from the same locus, are detected only rarely.
- PCR is an enzymatic reaction, therefore, the quality and concentration of template DNA, concentrations of PCR components, and the PCR cycling conditions may greatly influence the outcome. Thus, the RAPD technique is notoriously laboratory-dependent and needs carefully developed laboratory protocols to be reproducible.
- Mismatches between the primer and the template may result in the total absence of PCR product as well as in a merely decreased amount of the product. Thus, the RAPD results can be difficult to interpret.
- Lack of a prior knowledge on the identity of the amplification products.
- Problems with reproducibility (sensitive to changes in the quality of DNA, PCR components and PCR conditions).
- Problems of co-migration (do equal-sized bands correspond to the same homologous DNA fragment?). Gel electrophoresis can separate DNA quantitatively, cannot separate equal-sized fragments qualitatively (i.e. according to base sequence).

DEVELOPING LOCUS-SPECIFIC, CO-DOMINANT MARKERS FROM RAPDS

- The polymorphic RAPD marker band is isolated from the gel.
- It is amplified in the PCR reaction. The PCR product is cloned and sequenced.
- New longer and specific primers are designed for the DNA sequence, which is called the sequenced characterized amplified region marker (SCAR).

APPLICATIONS OF RAPD ANALYSIS

It has become widely used in the study of

- genetic diversity/polymorphism,
- germplasm characterization,
- genetic structure of populations,
- domestication,
- detection of somaclonal variation,
- cultivar identification,
- hybrid purity,
- genome mapping,
- developing genetic markers linked to a trait in question,
- population and evolutionary genetics,
- plant and animal breeding,
- animal-plant-microbe interactions,
- pesticide/herbicide resistance,

- animal behavior study, and
- forensic studies

RAPD markers exhibit reasonable speed, cost and efficiency compared with other methods; and RAPD can be done in a moderate laboratory. Therefore, despite its reproducibility problem, it will probably be important until better techniques are developed in terms of cost, time and labour.

ACKNOWLEDGEMENTS

The authors thank DBT, New Delhi (DBT/CTEP/01/201100105 dt. 23 March 2011), UGC, New Delhi (Reference No. F. No. 34-452/2008 (SR) dt. 30 Dec 2008), CSIR, New Delhi (No. 37(1362)/09/EMR-II dt. 3 March 2009) and DST, New Delhi (SR/FT/LS-148/2008 dt. 7 Oct 2009) for supporting research projects related to RAPD-PCR and DBT for the State Biotech Hub (BT/04/NE/2009 dt. 22nd Dec. 2010).

REFERENCES

1. Botstein D, White RL, Skolnick M & Davis RW (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Amer J Human Genet*, 32, 314-3310.
2. Erlich HA (1989). *PCR Technology Principles and Applications for DNA amplification*. Stockton Press, New York.
3. Williams JGK, Kubelik AR, Livak KJ, Rafalski JA & Tingey SV (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res*, 18, 6531-6535.
4. Welsh J & McClelland M (1990). Fingerprint genomes using PCR with arbitrary primers. *Nucleic Acids Res*, 18, 7213-7218.
5. Caetano-Annoles G, Bassam BJ & Gresshoff PM (1991). DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. *Biotechnology*, 9, 553-557.
6. Griffiths AJF, Miller JH, Suzuki DT, Lewontin RC & Gelbart WM (1996). *An Introduction to Genetic Analysis*, *6th edn*. W.H. Freeman and Co., NY.
7. Paran I & Michelmore RW (1993). Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Theor Appl Genetics*, 85, 985-993.
8. Bardakci F & Skibinski DOF (1999) A polymorphic SCAR-RAPD marker between species of *Tilapia*. *Anim Genet*, 30, 78-79.
9. Roholf FJ (1998). *NTSYS-pc Numerical Taxonomy and Multivariate Analysis System*, *Version 2. 00*. Exeter Software, Setauket, New York.
10. Sunnucks P (2000). Efficient genetic markers for population biology. *Trends Ecol Evol*, 15, 199-203.
11. Freeland JR (2005). *Molecular Ecology*. John Wiley and Sons, West Essex, UK.
12. Sambrook J, Fristch EF & Maniatis T (1989). *Molecular Cloning. A laboratory Manual*, *2nd edn*. Cold Spring Harbor, Cold Spring Harbor Laboratory Press, New York.
13. Innis MA, Gelfand DH, Sninsky JJ & White TJ (eds) (1990). *PCR Protocols: A Guide to Methods and Applications.* Academic Press, Inc., San Diego.
14. McPherson MJ, Quirke P & Taylor GR (eds) (1991). *PCR: A Practical Approach.* Oxford University Press, Oxford.