ScienSage

Journal of Advanced Scientific Research

Available online through http://www.sciensage.info

ISSN
0976-9595
Research Article

PHYLOGENIC STUDY TO FIND GENETIC DIVERSITY OF WILDLY GROWING *PHYLLANTHUS EMBLICA* IN VINDHYA REGION

Shashi Kumar Dubey*[1], Deepak Mishra[1], Azaz Khan[2], Megha Jha[2]

[1]*Department of Biotechnology, AKS University, Panna Road, Sherganj, Satna, Madhya Pradesh, India*

[2]*Pinnacle Biomedical Research Institute (PBRI), Near, Bharat Scout and Guides Campus, Shanti Marg, Shyamla Hills Road, Bhopal, Madhya Pradesh, India*

**Corresponding author: shashidubey86@gmail.com*

ABSTRACT

Correct genotype identification of medicinal plant material remains important for botanical drug industry. Limitations of chemical and morphological approaches for authentication have generated need for newer methods in quality control of botanicals. Aonla, the Indian Gooseberry (*Phyllanthus emblica* Linn, *P. emblica*) is an important minor fruit crop, widely grown in India due to its neutraceutical properties and having commercial significance. The present study was carried out to develop DNA based marker for identification of *P. emblica*. Investigations on the use of RAPD markers enabled us to estimate genetic variability among commercially cultivated 5 varieties. This study also enabled us to distinguish these varieties using a set of four decamer primers, which was otherwise difficult by using morphological markers. The genetic similarity was evaluated on the basis of presence or absence of bands. High degree of polymorphism was observed among the samples, suggesting the degree of genetic variability. The genetic distance was very close within the varieties. Cluster analysis revealed two different groups of varieties directly associated to their place of origin. RAPD markers were also able to differentiate varieties of same origin or even selection from same parents. This information can be used for identification of varieties and further crop improvement programmed.

Keywords: Genetic diversity, *Phyllanthus emblica,* Random amplification of polymorphic DNA (RAPD)

1. INTRODUCTION

Diversity in plant genetic resources provides opportunity for plant breeders to develop new and improved cultivars with desirable characteristics which include both farmer-preferred traits (yield potential and large seed, etc.) and breeders preferred traits (pest and disease resistance and photosensitivity, etc.). From the very beginning of agriculture, natural genetic variability has been exploited within crop species to meet subsistence food requirement and now it is being focused to surplus food for growing populations. In the middle of 1960s developing countries like India experienced the green revolution by meeting food demand with help of high-yielding and fertilizer responsive dwarf hybrids/varieties especially in wheat and rice. These prolonged activities that lead to the huge coverage of single genetic cultivars (boom) made situation again worse in other forms such as genetic erosion (loss of genetic diversity) and extinction of primitive and adaptive genes (loss of landraces) [1, 2]. The biodiversities belonging to any area are the source of economic wealth and pride of that area obliging it as a shining part of the national heritage. Each flora has a particular form of species that is specially adapted to a particular locus where it grows. For such species, differences between their morphological and molecular levels are identified and used for better management. Genetic diversity is the key pillar of biodiversity and diversity within species, between species and of ecosystems. However, the problem is that modern crop varieties, especially, have been developed primarily for high yielding potential under well endowed production conditions. Such varieties are often not suitable for low income farmers in marginal production environments as they are facing highly variable stress conditions [3]. Vindhya region is an extensive storehouse of the potential source of less known medicinal plants [4]. Delimitation of plant species is of central importance in many areas of biology, such as biogeography, ecology, population genetics, macro-evolution, phylogeny systematic, conservation biology and biodiversity [5, 6]. Theoretically, there is a strict relationship between a

genotype and a phenotype that means morphological and molecular makers must show an association but for most of the plant species, the efficiency of morphological and molecular markers for species delimitation may not be the same [7]. Correctly identifying the genotypes of the medicinal plant has eminence in the botanical drug industry [8]. *Phyllanthus* is a large genus comprises of over 800 species of the family Phyllanthaceae which is distributed in tropical and sub-tropical regions and *Phyllanthus emblica* (*P. emblica*, Euphorbiaceae) is accentuated among medicinal *Phyllanthus* species due to its effective pharmacological activity [9]. An official drug of Ayurvedic Pharmacopoeia [10] and Indian Herbal Pharmacopoeia [11], it forms a main ingredient of various multi-component formulations. The whole plant, *P. niruri* found in the Vindhya region of Eastern Uttar Pradesh, India is used to treat diuretic, jaundice [12]. A root of *P. niruri* tree belongs to family *Euphorbiaceae* is used to treat viral hepatitis. *P. emblica* is an important minor fruit crop having commercial significance and genetic diversity studies of *P. emblica* may assist in crop improvement for cultivable species [13]. Population genetic studies are performed with general goals to characterize the extent of genetic variation within species and account for this variation. The frequency of genes and the forces that affect their frequencies, such as migration, mutation, selection, and genetic drift are determinants of the amount of genetic variation within and between populations [14]. There is a scarcity of information about the global pattern of gene expression changes during the domestication of the wild species. Domestication or eventually cultivation of species affects the pattern of variation or shows decrease in gene expression and this decrease may have been associated with decreased genetic diversity [15]. Hence the genetic diversity study of *P. emblica* may bring some unknown facts related to the domestication of species. Medicinal properties of *Phyllanthus* being a center of attention, the species *P. emblica* found in the Vindhya region of MP requires attention for better management and needs classification on grounds of molecular properties. Also, there is a scarcity of data regarding genetic diversity related to the *P. emblica* found in MP we need to focus on this variety for further studies for the same. In our study, we have also included chemo profiling of the samples collected as the *P. emblica* is known for its medicinal importance. We collected five individual populations of *P. emblica* from the various habitats of the Vindhya region belonging to Madhya Pradesh (MP), India. Our study aims to find the genetic variability among the cultivated varieties using random amplification of polymorphic DNA (RAPD) analysis.

2. MATERIAL AND METHODS

2.1. Sample collection

The plant materials were collected, reported from Vindhya Region belonging to MP which included Sidhi and Rewa districts. The GPS location (Global Positioning System) data at each collection point were cataloged. All the varieties collected are mentioned in table 1. Fresh, young, disease-free leaves and fruits were collected for genetic diversity assessment using RAPD markers.

2.2. Methodology

We have performed the genomic analysis for molecular studies. For this purpose, genomic DNA was extracted from the sample using the standard protocol. Then isolated DNA was introduced to a polymerase chain reaction (PCR) using RAPD primers, after which electrophoresis of PCR product has been done and then DNA bands of interest and desired molecular weight were obtained on the gel.

2.2.1 DNA Isolation

Total genomic DNA of all the 10 samples (5 flowers and 5 leaves) were isolated using modified cetyl trimethyl ammonium bromide (CTAB) extraction method [16]. In brief, fresh leaf tissue (1-1.5 g) or dried flower powders (0.5 g) was ground in liquid nitrogen with polyvinyl pyrrolidone (PVP). Freshly prepared extraction buffer (containing 0.3% [v/v] b-mercaptoethanol) was added to adjust the pH to 7.5-8. The suspension was incubated at 65°C for 30 min with intermittent pH monitoring. After cooling at room temperature for 5 min an equal volume of chloroform- isoamyl alcohol (IAA) (24: 1) was added and centrifuged at 665 g for 10 min at 25 °C. DNA from aqueous layer was precipitated by adding 1/10th volumes of 3 M sodium acetate and 2 volumes of chilled ethanol at 4°C. The mixture was centrifuged at 665 g at 4 °C for 5 min. DNA pellet was washed with 70% (v/v) ethanol and dissolved in Tris-Cl-EDTA (pH 8.0) (TE) buffer after drying in SpeedVac (Savants, U.S.A.).

2.2.2 Quantification of extracted genomic DNA and Integrity Checking

Quantity of DNA was assessed using Nanodrop Spectrophotometer at the ratio of absorbance at 260 nm

and 280 nm such that pure DNA preparation is determined at a value of 1.8. A value less than 1.8 signifies presence of either proteins or phenol as impurities [17]. Quantity of DNA samples was calculated using the following formula:

Concentration of DNA= A260 X 50µg X dilution factor

Purity of DNA = A260: A280 ratio = A260/ A280.

2.2.3 PCR amplification

The PCR amplification was performed according to the method developed by McClelland *et al.* [18]. Decamer oligonucleotide primers RPI-3, RPI-8, RPI-9, and RPI-10 were designed for RAPD analysis of genomic DNA extracted from the samples. An aliquot stock solution of RAPD primers was prepared by adding 10 times Mili Q water of the nanogram concentration of RAPD primers to the primer stock to make a stock solution volume of 100 pmol. The working solution of primers was prepared by adding 10ml of primer stock solution and 90 ml of distilled water. The final working stock solution volume was 10 pmol. PCR reaction was carried out in 25.50µl reaction tubes, using 4 primers mentioned above. PCR reaction mixture contains PCR master mix (12.50µl), nuclease-free water (11.00 µl), Template DNA (01.00 µl), and RAPD primers (01.00 µl). Amplification was performed in a thermocycler. The standard cycle applied to amplify the DNA. The standard cycle consists of an initial denaturation step at 94°C for 5 min, followed by 10 cycles of 94°C for 45 sec, 35°C for 1 min, and 72°C for 1.5 min; followed by 40 cycles of 94°C for 45 sec, 35°C for 45 sec, and 72°C for 1 min; the final extension was held at 72°C for 10 min. Then PCR products were stored at 4ºC.

2.2.4 Qualitative analysis of extracted DNA

Qualitative analysis of amplifies DNA was done by agarose gel electrophoresis. In this method, the extracted DNA was allowed to run on 1% agarose gel. When the electrophoresis was over, the gel was transferred to the Gel Documentation System (Alpha innotech, USA). Visualized and obtained the DNA images using a 1.4-megapixel camera with autofocus, one transilluminator with 302 nm and 365 nm wavelength UV light source with high and low intensity, one white light table, and epi white facility. The molecular weight of DNA fingerprints was obtained with the help of software alpha view which is used for the analysis of Gel over which DNA has been run. The well-labeld images of the gels are shown in fig. 1-5.

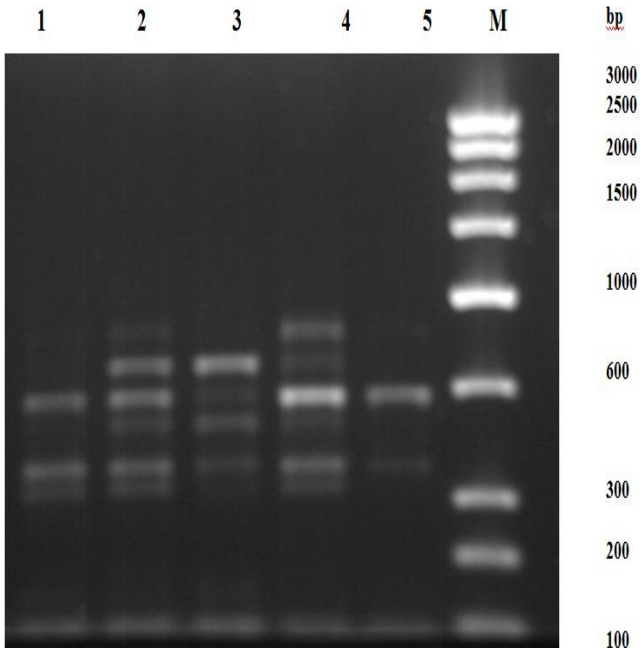

Fig. 1: RAPD fingerprints of five samples of *P. emblica* (Flower) obtained by RPI-3 primer. 1=S1F, 2=S2F, 3=S3F, 4=S4F and 5=S5F. Molecular weight marker=bp

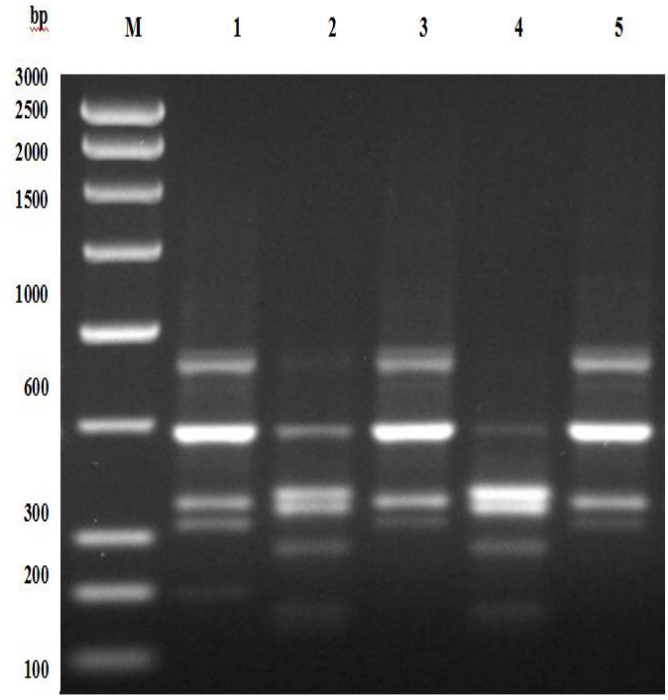

Fig. 2: RAPD fingerprints of five samples of *P. emblica* (Leaf) obtained by RPI-3 primer. 1=S1L, 2=S2L, 3=S3L, 4=S4L and 5=S5L. Molecular weight marker=bp

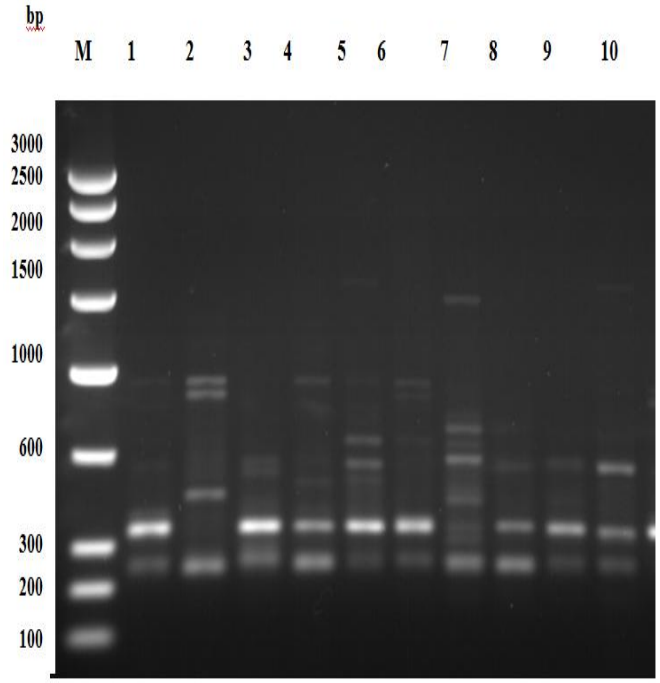

Fig. 3: RAPD fingerprints of five samples of *P. emblica* (Flower & Leaf) obtained by RPI-8 primer. 1=S1F, 2=S2F, 3=S3F, 4=S4F, 5=S5F, 6=S1L, 7=S2L, 8=S3L, 9=S4L and 10=S5L. Molecular weight marker=bp

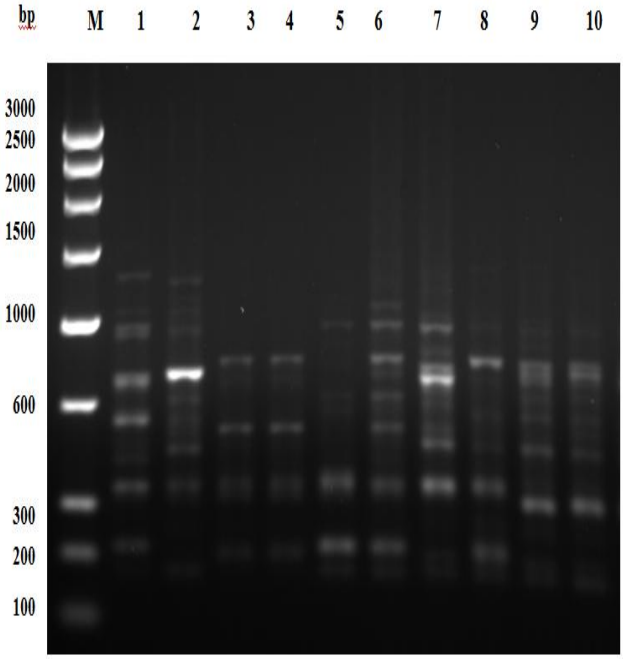

Fig. 4: RAPD fingerprints of five samples of *P. emblica P. emblica* (Flower & Leaf) obtained by RPI-9 primer. 1=S1F, 2=S2F, 3=S3F, 4=S4F, 5=S5F, 6=S1L, 7=S2L, 8=S3L, 9=S4L and 10=S5L. Molecular weight marker=bp

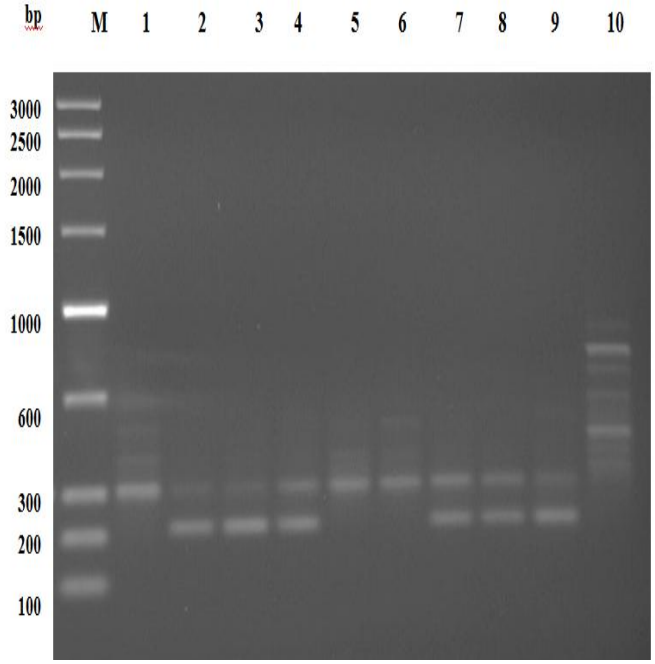

Fig. 5: RAPD fingerprints of five samples of *P. emblica* (Flower & Leaf) obtained by RPI-10 primer. 1=S1F, 2=S2F, 3=S3F, 4=S4F, 5=S5F, 6=S1L, 7=S2L, 8=S3L,9=S4L and 10=S5L. Molecular weight marker=bp

2.2.5 Data analysis

DNA banding patterns generated by RAPD are scored from the photograph as '1' for the presence of the band and '0' for its absence. Only clear bands were considered for the presence. The bands are considered polymorphic when they are absent in some samples. On the other hand, bands are considered monomorphic when they are present in all samples. Change in band intensity is not considered as polymorphism. A final RAPD data is generated and documented in the form of an MS Excel sheet which is used to calculate pair wise similarity co-efficient using the SIMQUAL format of the NTSYS-pc version 2.1 software packages. The similarity coefficients were noted in table 3 for all 10 samples. Cluster analysis was performed and the resulting similarity coefficients were used to construct dendrogram using the SAHN module of computer package program Numerical Taxonomy and Multivariate Analysis System (NTSYSpc). Amplified RAPD and ISSR products were analyzed by scoring polymorphic DNA bands and were compiled in a binary matrix in which 1 indicated the presence of band and 0 the absence of the band. The data were converted to distance matrices based on Nei and Li [19] unbiased

minimum distance. The distance matrices were then used to construct a dendrogram (1000 bootstrapping) by the unweighed pair-group method with arithmetic mean (UPGMA) using Tools for Population Genetic Analyses (TFPGAVer 1.3) [20].

3. RESULTS AND DISCUSSION

Five individual populations of *P. emblica* were collected from different areas (table 1) and subjected to RAPD analysis using Primer RPI-3, RPI-8, RPI-9, and RPI-10. After extraction and amplification of genomic DNA, the distinct and abundant RAPD fragments were recorded on gel. The total numbers of bands generated were 95 RAPD gel profiles. The purity check done by the spectrophotometer proved the integrity of the samples extracted (table 2).

Table 1: *P. emblica* plant species collected from different localities

S. No.	Sample Code	Area of study	Geographical location
1.	S1	Sirmour area (wardha ghat),Rewa	24.85°N 81.38°E
2.	S2	Semariya,Rewa	24°47′42″N 81°9′8″E
3.	S3	Rampur Naikin, Sidhi	24°20′23″N 81°28′29″E
4.	S4	Hanumana, Rewa	24°46′30″N 82°5′24″E
5.	S5	Mukundpur, Rewa	24.4218° N, 81.2436° E

Table 2: Spectrophotometric quantification of extracted genomic DNA

S. No.	Sample Code	OD 260/OD 280 ratio	Concentration(ng/µl)
1	S1F	1.91	850.2
2	S2F	1.90	851.1
3	S3F	1.86	752.4
4	S4F	1.78	723.2
5	S5F	1.69	654.4
6	S1L	1.64	631.2
7	S2L	1.72	723.4
8	S3L	1.79	759.1
9	S4L	1.68	645.2
10	S5L	1.74	726.3

Table 3: Similarity matrix among the samples of *P. emblica* collected from five locations of the Madhya Pradesh

	S1F	S2F	S3F	S4F	S5F	S1L	S2L	S3L	S4L	S5L
S1F	1									
S2F	1	1								
S3F	0.30769	0.30769	1							
S4F	0.46154	0.46154	0.45455	1						
S5F	0.23077	0.23077	0.18182	0.15385	1					
S1L	0.35714	0.35714	0.23077	0.63636	0.25	1				
S2L	0.4	0.4	0.5	0.53846	0.30769	0.33333	1			
S3L	0.42857	0.42857	0.30769	0.72727	0.23077	0.58333	0.61538	1		
S4L	0.14286	0.14286	0.18182	0.25	0.090909	0.15385	0.30769	0.33333	1	
S5L	0.17647	0.17647	0.13333	0.35714	0.14286	0.35714	0.16667	0.33333	0.33333	1

The sizes of the RAPD bands were placed in between 200-1500bp in length. The primer produced distinct polymorphic banding pattern in all the samples, the number of RAPD bands per primer was 23.5 as expected in *P. emblica*. The number of RAPD bands produced reveals Mendelian inherited characters, and number scoring revealed characters. In our study, the distinct banding patterns observed in *P. emblica* species

samples. The polymorphism was very high and RAPD values were useful to distinguish between the plant species, apparently diverse elements of species character. The similarity matrix was subjected to the Unweighted Pair Group Method for Arithmetic average analysis (UPGMA) and dendrogram obtained from the cumulative cluster analysis of four primer matrix using Jaccard's similarity coefficient delineated all 10samples of *P. emblica* (table 3 & fig. 6).

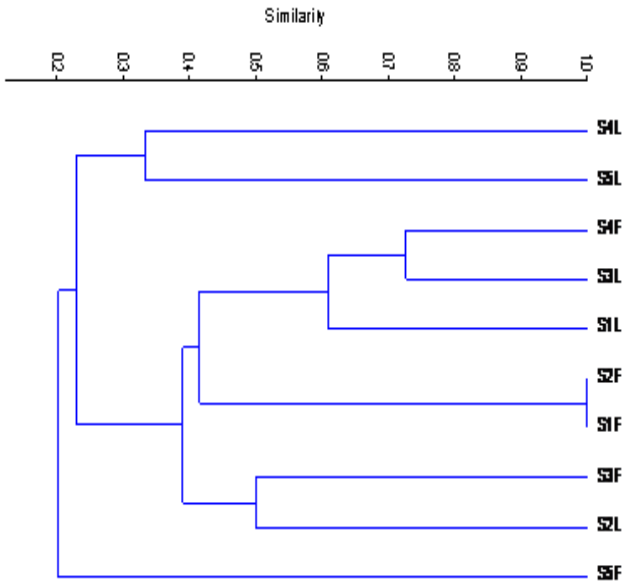

Fig. 6: Cluster analysis by RAPD fingerprinting among 10 samples of the *P. emblica* from five locations of the MP

Genetic relationship based on similarity matrix is used to measure the similarity for all marker type among populations [19] and the relationships between populations based on similarity index represented graphically in the form of a dendrogram is commonly called cluster analysis. The banding patterns generated by the five primers were scored in the form of a binomial matrix and were used to calculate the similarity index values of the DNA fingerprints according to Jaccard's coefficient [21]. Genetic distances (Euclidean distances) among the studied samples as revealed by morphometric markers were presented in table 4. The Jaccard's Distance matrix used to build a dendrogram that delineated all 10 samples of *P. emblica* (figure 7). The dendrogram in figure 6 was divided into 2 major clusters. Cluster I have only one accession S5F which depicts it uniqueness of genetic material or DNA and eventually is not closely related to other samples. Cluster II was subdivided into 2 sub-clusters; Sub clusters I have two accessions S4L and S5L which means they are closely related to each other and sub-clusters II has 7 accessions. The two most closely related accessions form sub cluster II was found to be S1F and S2F hence suggesting evidence of common source of origin. S4F, S3L and S1L are placed in same clade whereas S3F and S2L placed in same clade. Among 10 *P. emblica* species, the range of genetic similarity was from 2.44 to 3.873. A phylogenetic tree in fig.7 was created using the Neighbour-Joining method [22]. The tree consists of 3 major phylogenetic groups or clusters. Group or cluster I that includes S4F, S3L, S1L, S4L, and S5L that clearly shows close relatedness of S4F and S3L as they are placed in the same subgroup, group II with S1F, and S2F that are 100% close to each other and hence suggesting evidence of common source of origin, and the group III consisting of S3F and S2L.

Table 4: Jaccard's distance matrix among the samples of *P. emblica* collected from five locations of the Madhya Pradesh

0	S1F	S2F	S3F	S4F	S5F	S1L	S2L	S3L	S4L	S5L
S1F	0									
S2F	0	0								
S3F	3	3	0							
S4F	2.6458	2.6458	2.4495	0						
S5F	3.1623	3.1623	3	3.3166	0					
S1L	3	3	3.1623	2	3	0				
S2L	3	3	2.4495	2.4495	3	3.1623	0			
S3L	2.8284	2.8284	3	1.7321	3.1623	2.2361	2.2361	0		
S4L	3.4641	3.4641	3	3	3.1623	3.3166	3	2.8284	0	
S5L	3.7417	3.7417	3.6056	3	3.4641	3	3.873	3.1623	2.8284	0

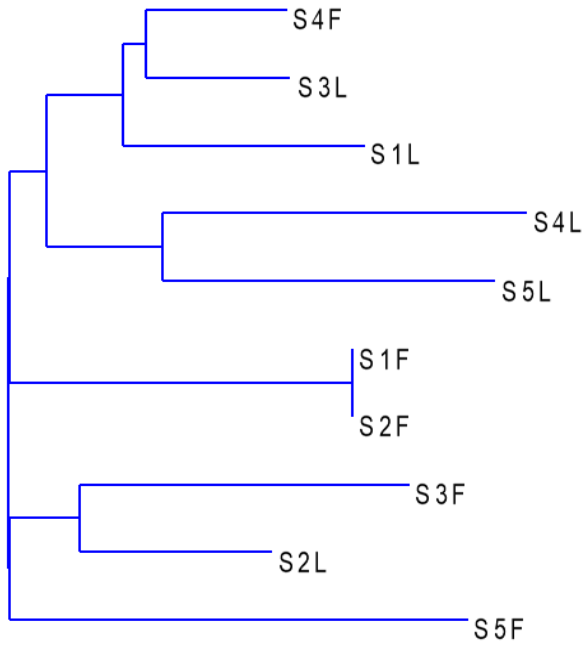

Fig. 7: Neighbour-Joining by RAPD fingerprinting among 10 samples of the *P. emblica* from five locations of the Madhya Pradesh

4. CONCLUSION

The term genetic diversity is referred to as the diversity present within different genotypes of the same species. We studied 5 variants and 10 samples of *P. emblica* and accessed their genetic constitution (genomic DNA) using RAPD. The total numbers of bands generated were 95 RAPD gel profiles. The number of RAPD bands per primer was 23.5 as expected in *P. emblica*. A phylogeny dendrograms created using UPGMA and Neighbor-joining methods revealed the genetic diversity among these samples of species. According to phylogeny analysis based on genomic DNA extraction and amplification data, these variants of species sample S1 and S2 from Sirmour and Semariya area of Rewa are closely related whereas S5 from Mukundpur, Rewa is most distinct or not closely related to other samples included in the study. All the 10 samples were delineated in dendrograms suggesting genetic diversity among *P. emblica* species forms the Vindhya region of MP.

5. REFERENCES

1. Narain P. *Cur. Sci.,* 2000: **79 (2)**:170-175.
2. Govindaraj M, Vetriventhan M, Srinivasan M. *Genet. Res. Int.,* 2015: 431487.
3. Evenson RE, Gollin D. *Science*, 2003; **300(5620)**:758-762.
4. Tiwari V, Soni P. *Int. J. Appl. Res. Tech.* 2016; **1(1)**:21-31.
5. Alfried P. Vogler RD. *Conserva. Biol.,* 1994; **8(2)**:354-363.
6. Rodrıguez JP, Balch JK, Rodrıguez-Clark KM. *Biodivers. Conserv.* 2007; **16**:183-209.
7. Duminil J, Michele MD. *Plant Biosyst.,* 2009; **143(3)**:528-542.
8. Warude D, Chavan P, Joshi K, Patwardhan B. *Biol. Pharm. Bull.,* 2006; **29(11)**:2313-2316.
9. Manissorn J, Sukrong S, Ruangrungsi N, Mizukami H. *Biol. Pharm. Bull.,* 2010; **33(10)**:1723-1727.
10. Government of India, the Ayurvedic Pharmacopoeia of India, Controller of Publications, Delhi, 2001.
11. Indian Herbal Pharmacopoeia, Indian Drug Manufacturer's Association, Mumbai, 2002.
12. Singh A, Singh P, Singh G, Pandey AK. *Int. J. Herbal Med.,* 2014; **2(2)**:31-37.
13. Singh SK, Meghwal PR, Pathak R. Proc. *Natl. Acad. Sci., India, Sect. B Biol. Sci.,* 2015; **85(4)**:971-978.
14. Gall K. Biotechnology: Concepts, Methodologies, Tools and Applications. IGI global, Hershey PA, USA. 1987; 693.
15. Liu W, Chen L, Zhang S, Hu F, Wang Z, Lyu J et. al. *BMC Evol. Biol.,* 2019; **19**:19.
16. Warude D, Chavan P, Joshi K, Patwardhan B. *Plant Mol. Biol. Rep.,* 2003; **21**:467.
17. Sambrook J, Russell D W. Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York, 2001.
18. McClelland M, Mathieu, Daude F, Welsh J. *Trends Gene.,* 1995; **11**: 242-246.
19. Nei M, Li WH. *Proc. Natl. Acad. Sci. USA,* 1979; **76(10)**:5269-5273.
20. Miller MP. Tools for Population Genetic Analyses (TFPGA) version 1.3. 1997; 1-33.
21. Jaccard P. Bull. Vaud. Soc. Nat. Sci., 1901; **37**: 241-272.
22. Saitou N, Nei M. *Mol. Biol. Evol.,* 1987; **4(4)**:406-425.