Development of a Statistical Uniqueness Index for DNA Fingerprinting Using Microsoft Excel

Lavanya Kumari Padherla[1],Amaravathi Yellari[2]

[1]Assistant Professor, Dept.of Statistics & Computer Applications, SMGR Agricultural College, ANGRAU, Udayagiri

[2]Scientist Pulses Scheme, Institute of Frontier Technology, Regional Agricultural Research Station, ANGRAU, Tirupati

***Corresponding Author**

DOI: https://doi.org/10.51244/IJRSI.2026.1307000319

Received: 06 August 2026; Accepted: 11 August 2026; Published: 18 August 2026

ABSTRACT

DNA fingerprinting of crop varieties is one of the most reliable methods for varietal identification and seed traceability in the seed industry. It provides authentic proof for breeders to prevent unauthorized use of varieties by private entities and to avoid infringement claims. It can be used as a molecular tool for genetic verification in seed certification. The technique also plays a key role in maintaining genetic integrity by preventing contamination and ensuring consistency in hybrid and open-pollinated varieties. In this way, DNA fingerprinting ensures the development of true-to-type seeds, which is vital for maintaining genetic purity in commercial cultivation and breeding programs. DNA fingerprint profiles of varieties and hybrids act as concrete evidence of genetic originality for legal protection through intellectual property rights and breeders' rights. In addition, DNA fingerprinting aids in the management of germplasm collections by enabling the formation of core and mini-core sets and identifying duplicates. Hence, the development of a Uniqueness Index calculator for molecular profiles helps to quantify the genetic distinctiveness of varieties quickly, suggesting the minimum number of markers required for marker profile development, facilitating the comparison of genetic profiles across varieties, and enabling the unambiguous identification of crop varieties.

Keywords: DNA Fingerprinting, Varietal identification, Seed trace, Genetic purity, Plant Breeders Rights, Germplasm and Uniqueness Index Calculator, MS-Excel

INTRODUCTION

The unambiguous identification of crop varieties is the most fundamental requirement for enforcing proprietary rights over plant varieties. It will facilitate the growth of the seed industry and the availability of high-quality seeds to farmers. At the global level, it has become necessary to register, characterize, and document varieties in the seed production chain. In India, a national system has been developed for the characterization of varieties based on DUS (Distinctness, Uniformity, and Stability). DUS testing forms the basis for a system of intellectual property protection for plant breeders known as Plant Breeders' Rights (PBR). Moreover, the “Protection of Plant Varieties and Farmers’ Rights Act (PVP&FRA)” was passed by the Parliament in 2001, which represented a significant leap forward. India ratified the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPs), which facilitates the adoption of an effective “sui generis” system for the protection of plant varieties. As precise genotype characterization is highly essential for the Plant Variety Protection (PVP) Act under the General Agreement on Tariffs and Trade (GATT), DNA fingerprinting has been formally admitted as a compulsory component of DUS testing since 2015. The National Research Centre on DNA Fingerprinting has been entrusted with the responsibility of fingerprinting the released varieties of nationally important crops since 2000 (Bhat, 2001).

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

Molecular profiling using DNA markers such as SSRs (Simple Sequence Repeats) or SNPs (Single Nucleotide Polymorphisms) provides a unique and stable genetic identity for each variety, which is crucial for distinguishing it from existing ones. This molecular characterization not only supports Distinctness, Uniformity, and Stability (DUS) testing but also strengthens the enforcement of Plant Breeders' Rights (PBR). DNA fingerprint profiles are non-tissue-specific, are minimally influenced by environmental factors, and follow simple Mendelian inheritance patterns. Along with morphological and agronomical data, DNA fingerprinting can establish the identity of genotypes without any ambiguity. The question underlying the use of DNA fingerprinting in cultivar identification is whether two or more cultivars have the same genetic fingerprint profile, or the probability that two or more cultivars will have the same pattern of DNA fragments. The molecular markers (M1, M2, and so on) are located at different positions in the genome and are inherited independently. Therefore, in theory, if the allele frequency of M1 in a population is 1 in 50, and that of M2 is 1 in 100, then only 1 in 5,000 cultivars would be expected to have both alleles at the respective markers. Generally, a greater number of markers are employed to generate a "unique DNA fingerprint profile," so that the probability of a false-positive match is extremely low (Rana and Bhat, 2017).

The test for uniqueness involves comparing the marker profile of the test variety with a local popular variety and a national check. If the profile shows no match or duplication with any existing entry, it is considered unique. The minimum number of DNA marker profiles required to establish the uniqueness of a DNA fingerprint in a crop variety is not universally fixed, as it largely depends on factors such as the crop species, the level of genetic diversity within the species, and the type of molecular markers employed. However, as per the guidelines issued by ICAR-NBPGR and the National Research Centre on DNA Fingerprinting (NRCDF), certain general standards have been established, particularly in the Indian context. For most crops, when using Simple Sequence Repeat (SSR) markers, a minimum of 10 to 15 highly polymorphic markers is typically considered sufficient to demonstrate distinctness and uniqueness. The number of required SNPs can vary widely, but usually, a panel of 30–50 informative SNPs are sufficient for most major crops, depending on the informativeness of the loci. The uniqueness test is scientifically validated using statistical parameters. To enhance the utility of DNA fingerprinting, the development of a statistical Uniqueness Index can provide a quantifiable measure of genetic distinctiveness, offering a robust framework for evaluating and comparing DNA profiles across different plant varieties.

MATERIALS AND METHODS

Molecular marker profiles developed by Amaravathi *et al.* (2014) in groundnut were used for the construction of the Uniqueness Index for DNA fingerprinting of twelve varieties released over a period of 25 years from the Regional Agricultural Research Station, Tirupati (Vasanthi *et al.*, 2012). The fingerprint profiles developed with SSR markers were converted into a binary 1/0 matrix, and the allelic profiles were superimposed to identify the common alleles shared among the varieties.

To obtain the level of confidence for identifying the 12 genotypes using 12 markers, the probability of an identical match by chance (Pi) was calculated following the method proposed by Wetton *et al.* (1987) and Ramakrishna *et al.* (1994). The stepwise procedure adopted was as follows:

Step 1: Scoring of marker data

The banding pattern generated by each of the 12 markers was scored for all 12 genotypes in binary form, where the presence of a band was recorded as 1 and the absence as 0.

Step 2: Pairwise comparison of genotypes

For Ggenotypes, the number of unique genotype pairs is $\text{Number of pairs} = \frac{G(G-1)}{2}$

For 12 genotypes, the number of unique genotype pairs is 66. All possible pairwise comparisons among the 12 genotypes were generated, resulting in 66 unique genotype pairs for further analysis.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

Step 3: Determination of total bands (N_A and N_B)

For each marker and for each pair of genotypes, the total number of bands present in genotype A (N_A) and genotype B (N_B) was counted as

$$N_A = \sum(\text{bands in genotype A})$$
$$N_B = \sum(\text{bands in genotype B})$$

Step 4: Identification of common bands (N_{AB})

The number of bands common to both genotypes (N_{AB}) for each marker was identified by comparing their banding profiles as $N_{AB}=\sum$(bands where both genotypes have 1)

Step 5: Calculation of similarity index (X)

The similarity index between the two genotypes for each marker was calculated using the Jaccard-like similarity formula:

$$X = \frac{2N_{AB}}{N_A + N_B}$$

which expresses the probability that a band present in one genotype is also present in the other.

Step 6: Estimation of average number of bands (n)

The average number of bands for each genotype pair was calculated as the simple mean of the total number of bands in the two genotypes

$$n = \frac{N_A + N_B}{2}$$

Step 7: Calculation of probability of identical match for each marker (Pi)

The probability of identical match by chance (P_i) for each genotype pair at a single marker was calculated under the assumption that the observed bands match independently. The probability of identical match for each marker was calculated as:

$$P_i = X^n$$

Where, X is the similarity index and n is the average number of bands.

Step 8: Computation of overall probability across 12 markers

The overall probability of identical match by chance for each genotype pair was obtained by multiplying the values across all 12 markers, assuming that the markers are independent. Thus, the probabilities across all markers were multiplied as follows, where $P_i$represents the probability of identical match by chance for the i^{th} marker:

$$P_{\text{overall}} = \prod_{i=1}^{m} P_i = P_1 \times P_2 \times P_3 \times \ldots \times P_{12}$$

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

Step 9: Construction of overall probability matrix

The overall probabilities for all genotype pairs (**$P_{overall}$** for all pairs) were arranged into a **12 × 12** matrix, where each cell represents the probability that the corresponding genotype pair would have an identical profile by

chance.

Step 10: Calculation of average and standard deviations of overall probability of identical match

The average overall probability of identical match by chance was calculated by summing the overall Pi values of all genotype pairs and dividing by the total number of pairs (66), providing a measure of the overall discriminatory power of the 12-marker system.

The average overall probability across all pairs (m =66 unique pairs):

$$\bar{P}_{\text{overall}} = \frac{\sum P_{\text{overall (all pairs)}}}{\text{number of pairs}} = \frac{\sum_{k=1}^{m} P_{\text{overall}}^{(k)}}{m}$$

Standard deviation of the overall probability:

$$SD = \sqrt{\frac{\sum_{k=1}^{m} (P_{\text{overall}}^{(k)} - \bar{P}_{\text{overall}})^2}{m-1}}$$

where m= total number of genotype pairs (66)

$P_{\text{overall}}^{(k)}$= overall probability of identical match by chance for the genotype pair;

$\bar{P}_{\text{overall}}$= mean overall probability of identical match by chance across all genotype pairs;

m = total number of genotype pairs (66).

Software used:

The computations were performed using **Microsoft Excel**, which provides a transparent, step-by-step approach for understanding and applying the proposed formulas. Unlike high-level statistical software or packages that may provide only final outputs, Excel allows researchers to **trace, verify, and understand each intermediate calculation**, thereby providing practical knowledge of **how and why the formulas work**. This facilitates independent application and validation of the methodology for different numbers of genotypes, markers, and alleles across various crops.

RESULTS AND DISCUSSION

The study comprised **12 genotypes (G1, G2, …, G12)** and **12 markers (M1, M2, …, M12)**. The following step-by-step procedure was followed to calculate the similarity indices for each pair of genotypes at each marker using Microsoft Excel.

Step1: The total number of amplified bands, irrespective of their base-pair size, obtained for each marker was recorded for all 12 genotypes, as shown in Table 1.

Table-1: Total no.of bands amplified for Marker:PM-375

Genotypes	100bp	105bp	116bp	**Total no.of bands**
G1	1	0	1	2

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

G2	0	1	1	2
G3	1	0	1	2
G4	1	0	1	2
G5	1	0	1	2
G6	1	0	1	2
G7	1	0	1	2
G8	1	0	1	2
G9	1	0	1	2
G10	0	1	1	2
G11	1	0	1	2
G12	1	0	1	2

Step:2: The number of common amplified bands, represented by **"1" in both genotypes**, was counted separately for each genotype pair at each marker.

Step-3: For each marker, the number of common amplified bands for all genotype pairs was listed, as shown in Tables 2–13.

Table-2: Common bands present in the pair of genotypes for Marker: PM-375

Total bands		2	2	2	2	2	2	2	2	2	2	2
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
2	G2	1										
2	G3	2	1									
2	G4	2	1	2								
2	G5	2	1	2	2							
2	G6	2	1	2	2	2						
2	G7	2	1	2	2	2	2					
2	G8	2	1	2	2	2	2	2				
2	G9	2	1	2	2	2	2	2	2			
2	G10	1	2	1	1	1	1	1	1	1		
2	G11	2	1	2	2	2	2	2	2	2	1	
2	G12	2	1	2	2	2	2	2	2	2	1	2

Table-3: Common bands present in the pair of genotypes for Marker pPGS seq 5D5

Total bands		1	3	1	1	3	2	2	2	2	2	3
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
3	G2	1										
1	G3	0	0									
1	G4	1	1	0								
3	G5	1	3	0	1							
2	G6	1	1	1	1	1						
2	G7	1	1	1	1	1	2					
2	G8	1	1	1	1	1	2	2				
2	G9	1	1	1	1	1	2	2	2			
2	G10	1	3	0	1	2	1	1	1	1		
3	G11	1	2	0	1	3	1	1	1	1	2	
2	G12	1	1	1	1	1	2	2	2	2	1	1

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

Table-4: Common bands present in the pair of genotypes for Marker Pm-3

Total bands		3	3	2	3	3	2	3	3	3	2	3
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
3	G2	3										
2	G3	2	2									
3	G4	3	3	2								
3	G5	3	3	2	3							
2	G6	2	2	1	2	2						
3	G7	3	3	2	3	3	2					
3	G8	3	3	2	3	3	2	3				
3	G9	3	3	2	3	3	2	3	3			
2	G10	2	2	1	2	2	2	2	2	2		
3	G11	3	3	2	3	3	2	3	3	3	2	
2	G12	2	2	1	2	2	1	2	2	2	1	2

Table-5: Common bands present in the pair of genotypes for Marker Pm-36

Total bands		2	2	3	2	2	2	2	2	2	2	2
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
2	G2	2										
3	G3	2	2									
2	G4	2	2	2								
2	G5	2	2	2	2							
2	G6	2	2	2	2	2						
2	G7	2	2	2	2	2	2					
2	G8	1	1	1	1	1	1	1				
2	G9	2	2	2	2	2	2	2	2			
2	G10	2	2	2	2	2	2	2	2	2		
2	G11	2	2	2	2	2	2	2	2	2	2	
2	G12	2	2	2	2	2	2	2	2	2	2	2

Table-6: Common bands present in the pair of genotypes for Marker PM-377

Total bands		4	4	3	4	3	3	3	3	3	3	3
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
4	G2	4										
3	G3	3	3									
4	G4	3	3	2								
3	G5	3	3	3	2							
3	G6	2	2	1	3	1						
3	G7	2	2	2	1	2	1					
3	G8	2	2	2	1	2	1	3				
3	G9	2	2	2	2	2	1	3	3			
3	G10	1	2	1	1	1	3	1	1	1		

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

3	G11	2	2	2	1	2	1	3	3	3	1	
3	G12	2	2	2	1	2	1	3	3	3	1	3

Table-7: Common bands present in the pair of genotypes for Marker PM-1489

Total bands		3	3	3	3	2	2	3	3	3	2	2
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
3	G2	3										
3	G3	3	3									
3	G4	3	3	3								
2	G5	2	2	2	2							
2	G6	1	1	1	1	1						
3	G7	3	3	3	3	2	1					
3	G8	3	3	3	3	2	1	3				
3	G9	3	3	3	3	2	1	3	3			
2	G10	2	2	2	2	2	1	2	2	2		
2	G11	2	2	2	2	2	1	2	2	2	2	
3	G12	3	3	3	3	2	1	3	3	3	2	2

Table-8: Common bands present in the pair of genotypes for Marker GM-384

Total bands		1	1	1	1	1	1	1	1	1	1	1
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
1	G2	0										
1	G3	0	1									
1	G4	1	0	0								
1	G5	1	0	0	1							
1	G6	0	1	1	0	0						
1	G7	0	1	1	0	0	1					
1	G8	1	0	0	1	1	0	0				
1	G9	1	0	0	1	1	0	0	1			
1	G10	0	1	1	0	0	1	1	0	0		
1	G11	1	0	0	1	1	0	0	1	1	0	
1	G12	1	0	0	1	1	0	0	1	1	0	1

Table-9: Common bands present in the pair of genotypes for Marker pPGS seq16g8

Total bands		1	2	2	2	2	0	1	1	2	1	2
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
2	G2	1										
2	G3	1	2									
2	G4	1	2	2								
2	G5	0	1	1	1							
0	G6	0	0	0	0	0						
1	G7	0	1	1	1	1	0					
1	G8	0	1	1	1	1	0	1				

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

2	G9	1	2	2	2	1	0	1	1			
1	G10	0	0	0	0	0	0	0	0	0		
2	G11	1	2	2	2	1	0	1	1	0	0	
0	G12	0	0	0	0	0	0	0	0	0	0	0

Table-10: Common bands present in the pair of genotypes for Marker IPAHM-23

Total bands		2	1	2	2	2	2	2	2	2	2	2
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
1	G2	0										
2	G3	0	1									
2	G4	2	0	0								
2	G5	2	0	0	2							
2	G6	0	1	2	0	0						
2	G7	2	0	0	2	2	0					
2	G8	2	0	0	2	2	0	2				
2	G9	2	0	1	2	2	0	2	2			
2	G10	0	1	2	0	0	2	0	0	0		
2	G11	2	0	0	2	2	0	2	2	2	0	
2	G12	2	0	0	2	2	0	2	2	2	0	2

Table-11: Common bands present in the pair of genotypes for TC1AO2

Total bands		5	4	5	4	4	5	5	4	4	4	4
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
4	G2	4										
5	G3	5	4									
4	G4	1	1	1								
4	G5	4	4	4	1							
5	G6	5	4	5	1	4						
5	G7	5	4	5	1	4	5					
4	G8	3	3	3	2	3	3	3				
4	G9	3	3	3	2	3	3	3	4			
4	G10	1	1	1	3	1	1	1	2	2		
4	G11	1	1	1	3	1	1	1	2	2	4	
5	G12	2	1	2	3	1	2	2	2	2	4	4

Table-12: Common bands present in the pair of genotypes for C2E11

Total bands		4	3	4	3	3	3	3	3	3	4	3
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
3	G2	3										
4	G3	4	3									
3	G4	3	3	3								
3	G5	3	3	3	3							
3	G6	3	3	3	3	3						

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

3	G7	3	3	3	3	3	3					
3	G8	3	3	3	3	3	3	3				
3	G9	3	3	3	3	3	3	3	3			
4	G10	4	3	4	3	3	3	3	3	3		
3	G11	3	3	3	3	3	3	3	3	3	3	
4	G12	4	3	4	3	3	3	3	3	3	4	3

Table-13: Common bands present in the pair of genotypes for TC5A06

Total bands		5	5	5	4	5	3	4	3	2	3	2
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
5	G2	4										
5	G3	5	4									
4	G4	3	4	3								
5	G5	5	4	5	3							
3	G6	3	3	3	3	3						
4	G7	4	3	3	2	4	2					
3	G8	3	3	3	2	3	2	3				
2	G9	2	2	2	2	2	2	2	2			
3	G10	3	3	3	2	3	2	3	3	2		
2	G11	2	2	2	1	2	1	2	2	1	2	
5	G12	4	4	4	3	4	2	4	3	2	3	2

Step-4: The data were entered in the Excel worksheet from cells **A1 to M14**. An empty table with the same headings was then created in cells **P1 to AB14** for marker **PM-375**, as shown in **Figure 1**.

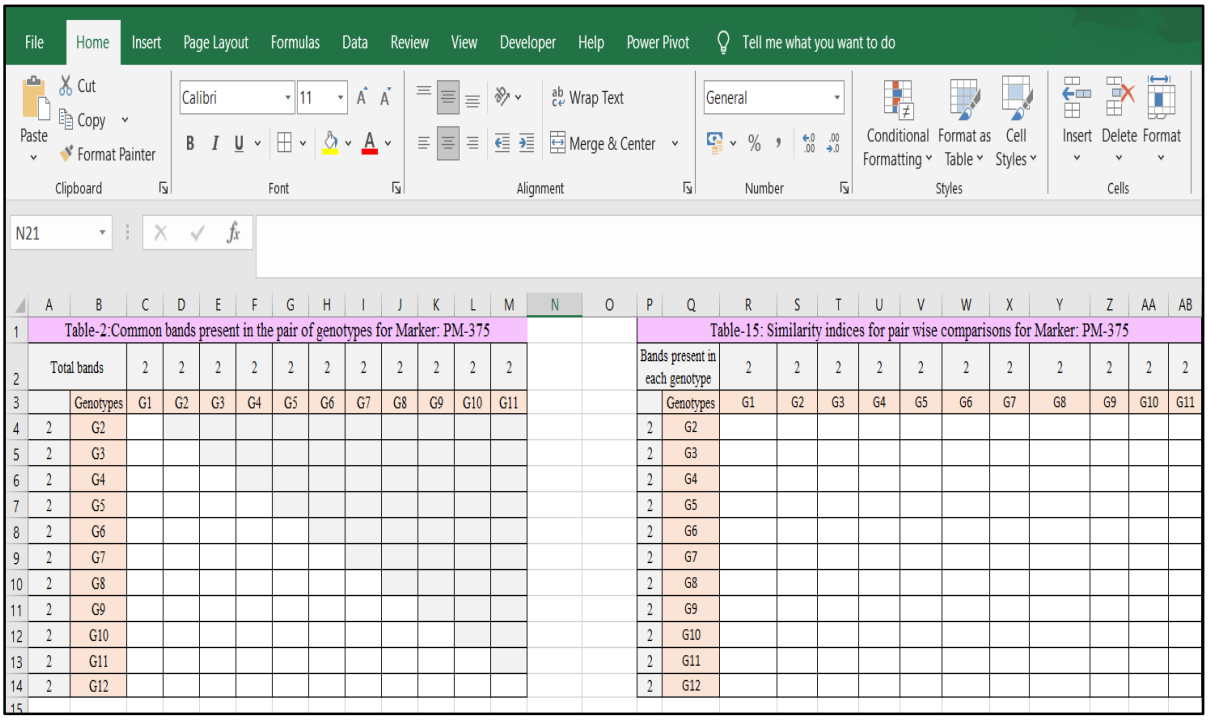

Figure-1: Input and output table entry in MS-Excel

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

Step-5: To compute the probabilities, place the cursor in cell R4 and enter the following formula: **=((2*C4)/(R$2+$P4))^(AVERAGE(R$2,$P4)).** Press the Enter key. The probability will then be displayed as 0.037 (Figure 2).

Evaluation of fragment patterns and statistical calculations were performed according to Wetton et al. (1987). A similarity coefficient (F)between two DNA fingerprints was calculated as $F=2N_{AB}/(N_A+N_B)$, where N_{AB} was the number of bands shared by fingerprints A and B, and N_A and N_B are the total number of bands present in fingerprints A and B, respectively. The probability that two different genotypes would exhibit identical fragment profiles by chance was then calculated as the average F raised to the mean number of fragments per genotype.

P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	A
Table-15: Similarity indices for pair wise comparisons for Marker: PM-375													
Bands present in each genotype		2	2	2	2	2	2	2	2	2	2	2	
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	
2	G2	0.250											
2	G3												
2	G4												
2	G5												
2	G6												
2	G7												
2	G8												
2	G9												
2	G10												
2	G11												
2	G12												

Figure-2: Computation of similarity index between G1 and G2

Step-6: Select cell **R4** and drag the formula downward from **R4 to R14**. Then, select cell **R5** and drag the formula horizontally from **R5 to S5**. Similarly, drag the formula from **R6 to T6, R7 to U7, R8 to V8, R9 to W9, R10 to X10, R11 to Y11, R12 to Z12, R13 to AA13, and R14 to AB14**. The respective probability values will then be automatically calculated and displayed in the corresponding cells, as shown in Figure-3.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

.

P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC
Table-15: Similarity indices for pair wise comparisons for Marker: PM-375													
Bands present in each genotype		2	2	2	2	2	2	2	2	2	2	2	
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	
2	G2	0.250											
2	G3	1.000	0.250										
2	G4	1.000	0.250	1.000									
2	G5	1.000	0.250	1.000	1.000								
2	G6	1.000	0.250	1.000	1.000	1.000							
2	G7	1.000	0.250	1.000	1.000	1.000	1.000						
2	G8	1.000	0.250	1.000	1.000	1.000	1.000	1.000					
2	G9	1.000	0.250	1.000	1.000	1.000	1.000	1.000	1.000				
2	G10	0.250	1.000	0.250	0.250	0.250	0.250	0.250	0.250	0.250			
2	G11	1.000	0.250	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.250		
2	G12	1.000	0.250	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.250	1.000	

Figure-3: Computation of similarity indices for all pairs of Genotypes

Step-7: Similarly, the probability for the remaining markers can be calculated using the respective tables (Tables 2 to 13), following the procedure described in Table 14 and Step 6. The computed similarity indices for each pair of genotypes across all the markers are presented in Table 15.

Table-14: Formula details for other markers for Probability of Identical match					
Input table. no	Range of input table	Range of probabilities table	Cell to place cursor	Formula to be typed in the cell for Probabilities	Output table.no.
2	A1 to M14	P1 to AB 14	R4	=((2*C4)/(R$2+$P4))^(AVERAGE(R$2,$P4))**(refer step.4)**	15
3	A16 to M29	P16 to AB 29	R19	=((2*C19)/(R$17+$P19))^(AVERAGE(R$17,$P19))	16
4	A31 to M44	P31 to AB 44	R34	=((2*C34)/(R$32+$P34))^(AVERAGE(R$32,$P34))	17
5	A46 to M59	P46 to AB 59	R49	=((2*C49)/(R$47+$P49))^(AVERAGE(R$47,$P49))	18
6	A61 to M74	P61 to AB 74	R64	=((2*C64)/(R$62+$P64))^(AVERAGE(R$62,$P64))	19
7	A76 to M89	P76 to AB 89	R79	=((2*C79)/(R$77+$P79))^(AVERAGE(R$77,$P79))	20
8	A91 to M104	P91 to AB 104	R94	=((2*C94)/(R$92+$P94))^(AVERAGE(R$92,$P94))	21
9	A106to M119	P106 to AB 119	R109	=((2*C109)/(R$107+$P109))^(AVERAGE(R$107,$P109))	22
10	A121 to M134	P121 to AB 134	R124	=((2*C124)/(R$122+$P124))^(AVERAGE(R$122,$P124))	23
11	A136 to M149	P136 to AB 149	R139	=((2*C139)/(R$137+$P139))^(AVERAGE(R$137,$P139))	24
12	A151 to M164	P151 to AB 164	R154	=((2*C154)/(R$152+$P154))^(AVERAGE(R$152,$P154))	25
13	A166 to M179	P166 to AB 179	R169	=((2*C169)/(R$167+$P169))^(AVERAGE(R$167,$P169))	26
		P181 to AB 194	R183	***=R4*R19*R34*R49*R64*R79*R94*R109*R124*R140*R155*R170***	**27**

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

Table-15: Probability of Identical match between two genotypes for Marker: PM-375											
Bands present in each genotype		2	2	2	2	2	2	2	2	2	2	2
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
2	G2	0.250										
2	G3	1.000	0.250									
2	G4	1.000	0.250	1.000								
2	G5	1.000	0.250	1.000	1.000							
2	G6	1.000	0.250	1.000	1.000	1.000						
2	G7	1.000	0.250	1.000	1.000	1.000	1.000					
2	G8	1.000	0.250	1.000	1.000	1.000	1.000	1.000				
2	G9	1.000	0.250	1.000	1.000	1.000	1.000	1.000	1.000			
2	G10	0.250	1.000	0.250	0.250	0.250	0.250	0.250	0.250	0.250		
2	G11	1.000	0.250	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.250	
2	G12	1.000	0.250	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.250	1.000

Table-16: Probability of Identical match between two genotypes for pPGS seq 5D5											
Bands present in each genotype		1	3	1	1	3	2	2	2	2	2	3
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
3	G2	0.250										
1	G3	0.000	0.000									
1	G4	1.000	0.250	0.000								
3	G5	0.250	1.000	0.000	0.250							
2	G6	0.544	0.101	0.544	0.544	0.101						
2	G7	0.544	0.101	0.544	0.544	0.101	1.000					
2	G8	0.544	0.101	0.544	0.544	0.101	1.000	1.000				
2	G9	0.544	0.101	0.544	0.544	0.101	1.000	1.000	1.000			
2	G10	0.544	1.577	0.000	0.544	0.572	0.250	0.250	0.250	0.250		
3	G11	0.250	0.296	0.000	0.250	1.000	0.101	0.101	0.101	0.101	0.572	
2	G12	0.544	0.101	0.544	0.544	0.101	1.000	1.000	1.000	1.000	0.250	0.101

Table-17: Probability of Identical match between two genotypes for Pm-3											
Bands present in each genotype		3	3	2	3	3	2	3	3	3	2	3
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
3	G2	1.000										
2	G3	0.572	0.572									
3	G4	1.000	1.000	0.572								
3	G5	1.000	1.000	0.572	1.000							
2	G6	0.572	0.572	0.250	0.572	0.572						
3	G7	1.000	1.000	0.572	1.000	1.000	0.572					
3	G8	1.000	1.000	0.572	1.000	1.000	0.572	1.000				
3	G9	1.000	1.000	0.572	1.000	1.000	0.572	1.000	1.000			
2	G10	0.572	0.572	0.250	0.572	0.572	1.000	0.572	0.572	0.572		

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

3	G11	1.000	1.000	0.572	1.000	1.000	0.572	1.000	1.000	1.000	0.572	
2	G12	0.572	0.572	0.250	0.572	0.572	0.250	0.572	0.572	0.572	0.250	0.572

Table-18: Probability of Identical match between two genotypes for Pm-36

Bands present in each genotype		2	2	3	2	2	2	2	2	2	2	2
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
2	G2	1.000										
3	G3	0.572	0.296									
2	G4	1.000	0.572	1.000								
2	G5	1.000	0.572	1.000	0.572							
2	G6	1.000	0.572	1.000	0.572	0.572						
2	G7	1.000	0.572	1.000	0.572	0.572	1.000					
2	G8	0.250	0.101	0.250	0.101	0.101	0.250	0.101				
2	G9	1.000	0.572	1.000	0.572	0.572	1.000	0.572	0.572			
2	G10	1.000	0.572	1.000	0.572	0.572	1.000	0.572	0.572	0.572		
2	G11	1.000	0.572	1.000	0.572	0.572	1.000	0.572	0.572	0.572	1.000	
2	G12	1.000	0.572	1.000	0.572	0.572	1.000	0.572	0.572	0.572	1.000	0.572

Table-19: Probability of Identical match between two genotypes PM-377

Bands present in each genotype		4	4	3	4	3	3	3	3	3	3	3
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
4	G2	1.000										
3	G3	0.583	0.583									
4	G4	0.316	0.316	0.141								
3	G5	0.583	0.583	1.000	0.141							
3	G6	0.141	0.141	0.037	0.583	0.037						
3	G7	0.141	0.141	0.296	0.012	0.296	0.037					
3	G8	0.141	0.141	0.296	0.012	0.296	0.037	1.000				
3	G9	0.141	0.141	0.296	0.141	0.296	0.037	1.000	1.000			
3	G10	0.012	0.141	0.037	0.012	0.037	1.000	0.037	0.037	0.037		
3	G11	0.141	0.141	0.296	0.012	0.296	0.037	1.000	1.000	1.000	0.037	
3	G12	0.141	0.141	0.296	0.012	0.296	0.037	1.000	1.000	1.000	0.037	1.000

Table-20: Probability of Identical match between two genotypes for PM-1489

Bands present in each genotype		3	3	3	3	2	2	3	3	3	2	2
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
3	G2	1.000										
3	G3	1.000	1.000									
3	G4	1.000	1.000	1.000								
2	G5	0.572	0.572	0.572	0.572							
2	G6	0.101	0.101	0.101	0.101	0.250						
3	G7	1.000	1.000	1.000	1.000	0.572	0.101					

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

3	G8	1.000	1.000	1.000	1.000	0.572	0.101	1.000				
3	G9	1.000	1.000	1.000	1.000	0.572	0.101	1.000	1.000			
2	G10	0.572	0.572	0.572	0.572	1.000	0.250	0.572	0.572	0.572		
2	G11	0.572	0.572	0.572	0.572	1.000	0.250	0.572	0.572	0.572	1.000	
3	G12	1.000	1.000	1.000	1.000	0.572	0.101	1.000	1.000	1.000	0.572	0.572

Table-21: Probability of Identical match between two genotypes for GM-384

Bands present in each genotype		1	1	1	1	1	1	1	1	1	1	1
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
1	G2	0.000										
1	G3	0.000	1.000									
1	G4	1.000	0.000	0.000								
1	G5	1.000	0.000	0.000	1.000							
1	G6	0.000	1.000	1.000	0.000	0.000						
1	G7	0.000	1.000	1.000	0.000	0.000	1.000					
1	G8	1.000	0.000	0.000	1.000	1.000	0.000	0.000				
1	G9	1.000	0.000	0.000	1.000	1.000	0.000	0.000	1.000			
1	G10	0.000	1.000	1.000	0.000	0.000	1.000	1.000	0.000	0.000		
1	G11	1.000	0.000	0.000	1.000	1.000	0.000	0.000	1.000	1.000	0.000	
1	G12	1.000	0.000	0.000	1.000	1.000	0.000	0.000	1.000	1.000	0.000	1.000

Table-22: Probability of Identical match between two genotypes for pPGS seq16g8

Bands present in each genotype		1	2	2	2	2	0	1	1	2	1	2
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
2	G2	0.544										
2	G3	0.544	1.000									
2	G4	0.544	1.000	1.000								
2	G5	0.000	0.250	0.250	0.250							
0	G6	0.000	0.000	0.000	0.000	0.000						
1	G7	0.000	0.544	0.544	0.544	0.544	0.000					
1	G8	0.000	0.544	0.544	0.544	0.544	0.000	1.000				
2	G9	0.544	1.000	1.000	1.000	0.250	0.000	0.544	0.544			
1	G10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		
2	G11	0.544	1.000	1.000	1.000	0.250	0.000	0.544	0.544	0.000	1.000	
0	G12	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table-23: Probability of Identical match between two genotypes for IPAHM-23

Bands present in each genotype		2	1	2	2	2	2	2	2	2	2	2
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
1	G2	0.000										
2	G3	0.000	0.544									

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

2	G4	1.000	0.000	0.000								
2	G5	1.000	0.000	0.000	1.000							
2	G6	0.000	0.544	1.000	0.000	0.000						
2	G7	1.000	0.000	0.000	1.000	1.000	0.000					
2	G8	1.000	0.000	0.000	1.000	1.000	0.000	1.000				
2	G9	1.000	0.000	0.250	1.000	1.000	0.000	1.000	1.000			
2	G10	0.000	0.544	1.000	0.000	0.000	1.000	0.000	0.000	0.000		
2	G11	1.000	0.000	0.000	1.000	1.000	0.000	1.000	1.000	1.000	0.000	
2	G12	1.000	0.000	0.000	1.000	1.000	0.000	1.000	1.000	1.000	0.000	1.000

Table-24: Probability of Identical match between two genotypes for TC1AO2

Bands present in each genotype		5	4	5	4	4	5	5	4	4	4	4
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
4	G2	0.589										
5	G3	1.000	0.589									
4	G4	0.001	0.004	0.001								
4	G5	0.589	1.000	0.589	0.004							
5	G6	1.000	0.589	1.000	0.001	0.589						
5	G7	1.000	0.589	1.000	0.001	0.589	1.000					
4	G8	0.161	0.316	0.161	0.063	0.316	0.161	0.161				
4	G9	0.161	0.316	0.161	0.063	0.316	0.161	0.161	1.000			
4	G10	0.001	0.004	0.001	0.316	0.004	0.001	0.001	0.063	0.063		
4	G11	0.001	0.004	0.001	0.316	0.004	0.001	0.001	0.063	0.063	1.000	
5	G12	0.010	0.001	0.010	0.161	0.001	0.010	0.010	0.026	0.026	0.589	0.589

Table-25: Probability of Identical match between two genotypes for TC2E11

Bands present in each genotype		4	3	4	3	3	3	3	3	3	4	3
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
3	G2	0.583										
4	G3	1.000	0.583									
3	G4	0.583	1.000	0.583								
3	G5	0.583	1.000	0.583	1.000							
3	G6	0.583	1.000	0.583	1.000	1.000						
3	G7	0.583	1.000	0.583	1.000	1.000	1.000					
3	G8	0.583	1.000	0.583	1.000	1.000	1.000	1.000				
3	G9	0.583	1.000	0.583	1.000	1.000	1.000	1.000	1.000			
4	G10	1.000	0.583	1.000	0.583	0.583	0.583	0.583	0.583	0.583		
3	G11	0.583	1.000	0.583	1.000	1.000	1.000	1.000	1.000	1.000	0.583	
4	G12	1.000	0.583	1.000	0.583	0.583	0.583	0.583	0.583	0.583	1.000	0.583

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

Table-26: Probability of Identical match between two genotypes for TC5A06												
Bands present in each genotype		5	5	5	4	5	3	4	3	2	3	2
	Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
5	G2	0.328										
5	G3	1.000	0.328									
4	G4	0.161	0.589	0.161								
5	G5	1.000	0.328	1.000	0.161							
3	G6	0.316	0.316	0.316	0.583	0.316						
4	G7	0.589	0.161	0.161	0.063	0.589	0.141					
3	G8	0.316	0.316	0.316	0.141	0.316	0.296	0.583				
2	G9	0.141	0.141	0.141	0.296	0.141	0.572	0.296	0.572			
3	G10	0.316	0.316	0.316	0.141	0.316	0.296	0.583	1.000	0.572		
2	G11	0.141	0.141	0.141	0.037	0.141	0.101	0.296	0.572	0.250	0.572	
5	G12	0.328	0.328	0.328	0.161	0.328	0.063	0.589	0.316	0.141	0.316	0.141

Table-27: Probability of Identical match between two genotypes for all 12 markers (P_overall)											
Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11
G2	0.0000000										
G3	0.0000000	0.0000000									
G4	0.0000186	0.0000000	0.0000000								
G5	0.0000000	0.0000000	0.0000000	0.0000018							
G6	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000						
G7	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000					
G8	0.0000000	0.0000000	0.0000000	0.0000033	0.0000946	0.0000000	0.0000000				
G9	0.0005543	0.0000000	0.0000000	0.0008139	0.0001096	0.0000000	0.0000000	0.1783664			
G10	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000		
G11	0.0000010	0.0000000	0.0000000	0.0000120	0.0000234	0.0000000	0.0000000	0.0006458	0.0000000	0.0000000	
G12	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
									Average		0.002737
									Sd		0.021952

Understanding the Table-27

- The rows and columns represent genotypes G1–G12.
- Each cell contains the probability (P) that the corresponding pair of genotypes has an identical profile across all 12 markers.
- Most values are extremely small or close to zero, which is expected because the probability of two unrelated genotypes having identical profiles across multiple markers is generally very low.
- The diagonal values are not presented (or may be considered as 1), since each genotype is identical to itself.

Key Observations

- The very small probabilities (approximately 0.0000) indicate that most genotype pairs have highly distinct marker profiles.
- The pair G8 and G9 shows the highest probability of an identical match (0.1783664). This relatively high value indicates that G8 and G9 share a greater degree of similarity in their marker profiles compared with the other genotype pairs.
- Some other genotype pairs show small but non-zero probabilities:

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

- G1 and G4: 0.0000186
- G4 and G5: 0.0000018
- G5 and G8: 0.0000946
- G4 and G11: 0.0000120

These values indicate minor similarities between the corresponding genotype pairs, but they are negligible compared with the probability observed for G8–G9.

Practical Implications

- **Unique Genotype**s: Most genotype pairs have near-zero probabilities, indicating that most genotypes can be distinguished based on their marker profiles.
- **Potentially Similar Genotypes**: G8 and G9 have a relatively high probability (0.178), suggesting that they may be closely related or share a greater degree of similarity in their marker profiles. However, this probability alone does not confirm that they are duplicates or share a common lineage.
- **For Breeding or Diversity Studies**: Genotype pairs with probabilities greater than 0.01 may be considered for further investigation of potential genetic similarity or redundancy. In the present dataset, only the G8–G9 pair exceeds this threshold and therefore warrants further examination.

Significant findings of table-27:

Table 27 represents the overall probability of an identical match (P) between pairs of 12 genotypes across 12 markers, reflecting the likelihood that any two genotypes share the same alleles at all marker loci. Most values in the table are extremely close to zero, indicating that the majority of genotypes are genetically distinct and can be reliably differentiated using these markers. Only a few pairs show small, non-zero probabilities, such as G4 with G1 (0.0000186), G5 with G4 (0.0000018), and G8 with G5 (0.0000946), suggesting minimal allele sharing that is likely due to chance rather than close relatedness.

A notable exception is the pair G8 and G9, which have a relatively high match probability of 0.1783664, indicating significant genetic similarity between these two genotypes. This elevated probability may suggest a potential duplication, shared ancestry, or close relationship, and warrant further investigation if maintain unique genetic identities is important. Overall, the data suggests that the set of 12 markers provides strong discrimination among the genotypes, with most pairs being highly distinct, while only one pair exhibits a considerable probability of matching. Further, the average and standard deviation of the overall probability of an identical match (P_overall) across the 12 markers were calculated for each pair of genotypes.

Step-8: The average and standard deviation of the overall probability of an identical match by chance were calculated using:

$\bar{X}_{P\_overall}$ = AVERAGE (R183:AB193) and

$SD_{P\_overall}$ =STDEV(R183:AB193) as shown in Table 27 below.

Average of overall probability of identical match by chance = 0.002737 (≈ 0.003)

Standard deviation of overall probability of identical match by chance = 0.021952 (≈ 0.022)

As the average probability of an identical match for each pair of genotypes across all 12 markers (Average of P_overall) is 0.003, this means that, on average, the probability that any two genotypes have identical DNA fingerprints across all 12 markers purely by chance is 0.3%. It is an extremely low value, indicating that the markers are highly discriminating, and that such matches are almost certainly due to true genetic similarity rather than random chance. At the same time, the SD is 0.022. It can be observed that the SD is larger than the mean, which suggests considerable variation among genotype pairs. Some pairs may have slightly higher probabilities of chance matches (up to approximately 2.2%), but most are still well below 5%.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

CONCLUSION

The present study, based on data generated from 12 molecular markers applied to 12 groundnut varieties developed and maintained at the Regional Agricultural Research Station, Tirupati, functioning under Acharya N. G. Ranga Agricultural University, Andhra Pradesh, demonstrated that the use of these markers provided a high level of confidence in discriminating the genotypes through the estimation of the probability of identical match by chance (P_i). The step-wise calculation of similarity indices and multilocus P_i values, following the approaches of Wetton et al. (1987) and Ramakrishna et al. (1994), revealed that most genotype pairs exhibited extremely low probabilities of identical match, indicating the strong discriminatory power of the marker set. The very small P_i values observed for the majority of comparisons confirm that the likelihood of two unrelated genotypes sharing identical marker profiles by chance is negligible. Only a few genotype pairs showed relatively higher Pi values, suggesting closer genetic similarity or possible relatedness, which may be attributed to shared ancestry or common breeding backgrounds. Overall, the results validate the effectiveness of the selected markers for reliable genotype identification and fingerprinting. The Excel-based computational programme developed and used in this study enables simple, transparent, and reproducible estimation of the probability of identical match and can be effectively used for routine varietal identification, germplasm characterization, seed certification, and plant variety protection programmes.

REFERENCES

1. Amaravathi, Y., Vasanthi, R.P., Siva Kumar, E., Purushotham M and Giridhara Krishna T. 2014. DNA fingerprinting of groundnut Arachis hypogaea L. varieties of Tirupati using SSR markers. Electronic Journal of Plant Breeding, 5 4, 677-687.
2. Bhat K.V. 2001. DNA Fingerprinting and cultivar identification. National Research Centre on DNA finger printing, NBPGR, New Delhi-110012
3. Lynch,M.(1988). Estimation of relatedness by DNA fingerprinting. Molecular Biology and Evolution, 5, 584–599.
4. Powell, W., Machray, G. C., & Provan, J. (1996). Polymorphism revealed by simple sequence repeats. Trends in Plant Science, 1, 215–222.
5. PPV&FR Authority (2009). Guidelines for the Conduct of Test for Distinctiveness, Uniformity and Stability.
6. Ramakrishna, W., Lagu, M. D., Gupta, V. S., & Ranjekar, P. K. (1994).DNA fingerprinting in rice using oligonucleotide probes specific for simple repetitive DNA sequences. Theoretical and Applied Genetics, 88, 402–406.
7. Rana M.K. and Bhat K.V. (2017) DNA Fingerprinting in Plants: Standard Operating Methods and Protocols, ICAR-National Bureau of Plant Genetic Resources, New Delhi, India, 73 pp.
8. Singh, R.K., et al. (2013). DNA Fingerprinting in Crop Plants. Indian Journal of Genetics and Plant Breeding, 73(1): 1–9.
9. Sundaram, R.M., et al. (2008). DNA fingerprinting and its applications in crop improvement. Current Science, 94(5): 612–619.
10. Tessier, C., David, J., This, P., Boursiquot, J. M., & Charrier, A. (1999). Optimization of the choice of molecular markers for varietal identification in Vitis vinifera L. Theoretical and Applied Genetics, 98, 171–177.
11. Vasanthi, R.P., John, K., Mohan Vishnuvardhan, K., Purushotham, M., Sirisha, K., Thejaswi, Y. And Giridhar Krishna T. 2012. In: Groundnut varieties developed/ Released by RARS, Tirupati (AP). Pp.17-26.
12. Wetton, J. H., Carter, R. E., Parkin, D. T., & Walters, D. (1987). Demographic study of a wild house sparrow population by DNA fingerprinting. Nature, 327, 147–149.