

Genetic Diversity Assessment of *Bixa orellana* L., a Natural Seed Dye Species of the Tropics, using ISSR Markers

Manasa Mogilicharla

SKLTSHU/Forest College and Research Institute

Reeja S

reejafcrits@gmail.com

SKLTSHU/Forest College and Research Institute

VBS Lalitha T

SKLTSHU/Forest College and Research Institute

Arjun Ramachandran

SKLTSHU/Forest College and Research Institute

Shalini Mudalkar

SKLTSHU/Forest College and Research Institute

Kumaran K

Anbil Dharmalingam Agricultural College and Research Institute, TamilNadu Agricultural University

Sreedhar Bodiga

SKLTSHU/Forest College and Research Institute

Chiranjeeva Milkuri

SKLTSHU/Forest College and Research Institute

Research Article

Keywords: *Bixa orellana*, PIC, unique bands, genetic diversity, UPGMA, ISSR markers

Posted Date: June 5th, 2025

DOI: <https://doi.org/10.21203/rs.3.rs-6660535/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Bixa orellana.L is an important dye yielding tree. The bixin content in the seeds yields a red colour pigment in varying colours which is used as dye. The variation in bixin content could be attributed to both environmental and genetic changes. In this study a total of 24 genotypes of *Bixa orellana*.L were raised from FC&RI, Annatto gene bank of Mettupalayam, India and their genetic diversity were analysed using ISSR markers. Out of 33 primers used, 23 primers were effectively screened for amplification and their polymorphism ranged from 50–100% indicating high genetic diversity among the genotypes. The values of polymorphic information content (PIC) ranged from 0.50–0.60. The other marker traits viz., effective marker ratio, discriminating power and heterozygosity are in higher values revealing the effectiveness of the ISSR markers. The total number of bands varied from 6 to 13. Among all the bands generated, 22 unique bands were detected in eight genotypes, out of which TSMLGB18 produced maximum number of unique bands. The UPGMA clustering pattern revealed that 24 genotypes were divided into seven distinct clusters. The genetic relationship obtained through Jaccard's similarity coefficient ranged between 0.57 to 0.88. The result provided valid guidelines for collection, conservation and characterization of *Bixa orellana* genetic resources.

Introduction

Bixa orellana is a medicinal dye source plant. It is normally admitted as “annatto”, “sindhur”, “lipstick tree”. Originally, it is a native of South American region and has its place in their culinary as a food ingredient and the most important ingredient of beauty products (Alcazar-Alay *et al.*, 2017). The seed coat extract gives out red orange pigment dye known as bixin that industrially competes well in the food and textile industries as a natural colorant and dye. It is the second most commercial non- toxic food colorant as compared to all other natural food colorants (Lopa pattanaik *et al.*, 2019). Other than the fact that it is used in textile industries, it has medicinal values that escalates its demands. It is a small tree or shrub that is perennially foliated and belongs to Bixaceae family. It is the only one in the Bixaceae family that has well-known uses. It possesses wide adaptability and is well distributed all over India in the following states Tamilnadu, Kerala, Karnataka, Chhattisgarh, Maharashtra, Madhya Pradesh, Gujarat and Andhra Pradesh (Kumaran K,2014). It may flourish in the temperature between 24°C – 38°C with various soils. A mean annual rainfall demand was 800-1500mm. It can even reach the height of 6 meters, if it will be provided with good irrigation (Kumaran K,2014).

The leaf is simple, smooth when mature, long petioles and are arranged alternating the branches (Franco *et al.*, 2008; Leal and Clavijo, 2012).

The flowers are vibrant and big, color ranges from white through to shades of pink and violet. They are bisexual and possess five sepals and they will usually see them flower right at the tips of the branches. The fruit is dehiscent, ovoid capsule made up of two/three carpels; covered by soft spines when young. It can be reddish, greenish or brownish and about 3–5 cm in length; it contains dark seeds covered in a reddish waxy Aril (bixin) with an unsettling smell. The fruits are borne in clusters as every fruit can hold

up to 17 pods (Dequigiovani *et al.*, 2017). The seeds in a given fruit may vary from 50–60 seeds. It is capable of adjusting itself to numerous ecological and edaphic conditions (Kumaran K *et al.*, 2014), means that, there is enough genetic variability available which can be exploited for its appropriate repercussions. Environmental factors have a dominant effect on the presence of bixin in seeds (Smith *et al.*, 2006). Due to environmental impact amount of bixin varies from place to place thereby suggesting variation is on population basis and not on individual variation (Vilares *et al.*, 1992).

The study of variation by the help of morphometric techniques was the dominating one earlier, and now, it is refined by the help of molecular techniques to evaluate the genetic variation which is less made in *Bixa orellana*. Environmental influences have a minimal impact on molecular markers and one can carry out analysis from DNA at any stage of development of the organism. PCR-assisted molecular markers – like random amplified polymorphic DNA (RAPD; Methods, such as, inter-simple sequence repeat (ISSR; Zietkiewicz *et al.*, 1994), amplified fragment length polymorphism (AFLP; Vos *et al.*, 1995), sequence-related amplified polymorphism (SRAP; Li & Quiros 2001)—have emerged. However, little work is known of the diversity and genetic structure of this species (Diquigiovanni *et al.*, 2017). Until recently the genetic diversity analysis on that of DNA markers viz. microsatellites, SRAP and RAPD (Sequence Related Amplified product and Randomly Amplified Polymorphic DNA) were used to evaluate genetic diversity in *bixa*. The current research attempts to evaluate the reliability of ISSR markers in the task of studying genetic variability of *Bixa orellana* which wasn't realized earlier. ISSR comes to the dominant marker class; they are emerged as an alternative system, with reliability and benefits of SSR markers and therefore it is reliable, fast and very informative system for studying the diversity (Zhang *et al.*, 2017). The present study made an attempt to study genetic diversity in the studied genotypes of *Bixa orellana*.

Materials and Methodology

Plant material;

To assess the genetic variation, seeds of *Bixa* were brought from field germplasm bank of Forest college & Research Institute (FCRI), (latitude 11° 9'N and longitude 76° 56'E) TNAU, Mettupalayam. The seeds were the individual collections from the progenies which are initially raised from the seeds of plus trees obtained from different parts of Southern India. For the purpose of this study, seeds from 24 such progenies (genotypes) were used to raise the seedlings & assembled at FCRI, Mulugu (17° 72'54.4" N to 78° 63'29.6" E), Telangana and were given accession codes (Table 1). Genomic DNA was extracted from young leaves following the standard method of CTAB, based on the procedure given by Doyle and Doyle (1991) with some minor changes.

Table 1
Details of accession number
and genotypes of *Bixa*
orellana.L

Accession No.	Genotypes
TSMLGB1	G1
TSMLGB2	G2
TSMLGB3	G3
TSMLGB4	G4
TSMLGB5	G5
TSMLGB6	G6
TSMLGB7	G7
TSMLGB8	G8
TSMLGB9	G9
TSMLGB10	G10
TSMLGB11	G11
TSMLGB12	G12
TSMLGB13	G13
TSMLGB14	G14
TSMLGB15	G15
TSMLGB16	G16
TSMLGB17	G17
TSMLGB18	G18
TSMLGB19	G19
TSMLGB20	G20
TSMLGB21	G21
TSMLGB22	G22
TSMLGB23	G23
TSMLGB24	G24

The leaf tissues of *Bixa orellana* (400mg) were grinded into to fine powder in liquid nitrogen and homogenized it with extraction buffer by adding 3ml of CTAB buffer,1%PVP, 250 µl of beta-mercapto-

ethanol and incubated for 30min at By using the ratio of 24:1 v/v of chloroform and isoamyl-alcohol solution supernant was extracted twice. 25 µl of RNase A was introduced in the extract which was then kept for incubation at room temperature. With the chilled iso-propanol, the DNA was pelleted and washed 2 times with 70% ethanol. Then air dried pellet added with 200 µl of nuclease free grade water and stored at -20°C. The quality of DNA was tested in nanodrop (ND1000, Thermo Scientific, US) spectrophotometer at 260 and 280 nm and quantity was measured in ng/µl. Portions that had between 200 µl to 600 µl DNA were best suited for PCR process.

ISSR amplification

Total of 33 random UBC primers were screened initially to assess genetic diversity among 24 genotypes of *B.orellana*. The 10 µl volume of ISSR PCR reaction mixture composed of 5 µl PCR emerald master mix, 2 µl water, 1.5 µl selected primer, 1.5 µl DNA sample. PCR reactions were performed as, i) denaturation was done at 94° for 4 minutes followed by ii) Annealing of primer at which temperature was specific to each primer and run for 1 minute then iii) Extension was done at 72°C for 2minutes. The reaction was repeated for 30 cycles. The amplified product was run for 1 hour at 100 V in 1.2% agarose gel (Himedia) immersed in 1X TBE buffer by placing the 1kb DNA ladder as reference. After running, the gel was scanned under UV Gel scanner and photographed in the system (Biorad).

Data analysis

DNA fingerprint patterns received were transformed into binary data matrices consisting of arrays of 0 and 1 s. The bands were awarded points based on their presence (1) or lack (0) of bands of different molecular weight sizes. Polymorphic and reproducible bands were the only ones taken into account for the analysis. Data were analysed with SIMQUAL route for creating Jaccard's similarity coefficient in NTSYS-pc version 2.02i (Numerical Taxonomy System) (Rohlf 1998). Dendrograms were done with using similarity matrices UPGMA (Unweighted Pair Group Method with Arithmetic Average) algorithm. The percentage polymorphism was done in Microsoft-excel. Some of the important measures which states the informativeness of marker was calculated using an online tool i.e., marker efficiency calculator (Amiryousefi et al ,2018;<https://irscope.shinyapps.io/iMEC/>). Polymorphic Information Content (PIC) and Heterozygosity(H) were calculated using Botstein et al. (1980) and Liu (1998).

$$PIC = 1 - (\sum p_i^2 - \sum p_i^2 p_j^2)$$

When p_i and p_j are population frequency of the i th and j th allele.

$H = 1 - \sum p_i^2$, in this p_i is the frequency of the allele for the i th allele.

Effective Multiplex Ratio (EMR) and Marker Index (MI) was calculated according to Powell *et al.* (1996).

$EMR = n \beta$, $\beta = np / (np + nnp)$, p and np represent polymorphic and nonpolymorphic fraction of the markers, hence np and nnp refer to the counting numbers of the same.

$MI = EMR * H_{avp}$,

H_{avp} refers to an average heterozygosity for the polymorphic makers. It is equal to $\sum H_n / np$, where, H_n represents the markers' polymorphic fraction's heterozygosity, and the summation is done over all the polymorphic loci np .

The discriminating power of a marker was determined by $D = 1 - C$ (Tessier et al., 1999).

Where, $C = \sum p_i (Np_i - 1) / (N - 1)$,

N – Individuals, C – Totality of loci for all the patterns formed by the primer.

Results

Among the 33 ISSR primers used, 23 primers that yielded distinct, reproducible and polymorphic banding profiles were screened for subsequent molecular characterization (Table 2). The total number of bands generated is 249 and per primer it is 10.8. Maximum number of bands were generated by UBC 826(9) while minimum was reported by ISSR-O4(6) with an average band size varied from 100bp to 3500bp. The polymorphism ranged from 50–100% among all genotypes. Of the 23 primers used in the study, 15 primers have shown 100% polymorphism viz., UBC841, UBC857, UBC826, UBC815, UBC888, UBC891, UBC810, UBC845, UBC847, UBC864, UBC880, ISSR-O3, ISSR-O4, ISSR-O6, ISSR-F2 (Fig. 1). Minimum polymorphism was reported in UBC823 (50%) followed by UBC 864(62.5%) (Fig. 2). The extent of polymorphism by all primers was presented in Table 3. Among all the bands generated, twenty-two bands were detected as unique among eight genotypes (Table 4), out of which TSMLGB18 was recorded for maximum number of unique bands (Fig. 3&Figure 4). Molecular markers particular to 24 genotypes were identified; this can be transformed into SCAR primers, to increase accuracy and reliability of the marker assisted selection. Table 5 were presented to show the molecular marker characters. The highest PIC was showed by UBC 809(0.60) followed by ISSR O6(0.59) and the minimum was reported for the UBC 857(0.505). All the 23 markers had > 0.50 PIC reporting the informative nature of the marker. The heterozygosity ranged from 0.2 to 0.6. In UBC 864 the highest values for PIC was 0.60 and that of heterozygosity was 0.66. The estimates of marker properties for ISSRs had wide range. The effective MP ratio is 1 for all the primers. Marker index (MI) was 0.66(UBC 864) to 0.24(ISSR O6). The maximum discriminating value was found in the case of UBC811 (0.86) and minimal in UBC 809(0.31).

Table 2
Primers selected for amplification and their specific annealing temperatures

S.NO	Primer	Sequence	Annealing Temperature
1	UBC 811	GAGAGAGAGAGAGAGC	43°C
2	UBC 841	GAGAGAGAGAGAGAYC	40°C
3	UBC 857	ACACACACACACACACG	59°C
4	UBC 826	ACACACACACACACACC	54°C
5	UBC 815	CTCTCTCTTCTCTCTG	41°C
6	UBC 888	BDBCACACACACACACA	44°C
7	UBC 889	DBDACACACACACACAC	39°C
8	UBC 891	HVHTGTGTGTGTGTGTG	52°C
9	UBC 813	CTCTCTCTCTCTCTCTT	41°C
10	ISSR M5	(GA)9-C	43°C
11	ISSR-03	(CT)8-GC	51°C
12	ISSR-04	(CA)6-AC	34°C
13	ISSR-06	(CT)6-AG	34°C
14	ISLA(CT)8TG	(CT)8-TG	43°C
15	ISSR F2	(AG)8-CG	54°C
16	UBC-809	AGAGAGAGAGAGAGAG	40°C
17	UBC-810	GAGAGAGAGAGAGAGT	44°C
18	UBC-823	TCTCTCTCTCTCTCTCC	43°C
19	UBC-840	GAGAGAGAGAGAGAGT	51°C
20	UBC-845	CTCTCTCTCTCTCTCTRG	43°C
21	UBC-847	CACACACACACACACC	44°C
22	UBC-864	ATGATGATGATGATGTG	49°C
23	UBC-880	GGAGAGGAGAGGAGA	49°C

Table 3
Percentage polymorphism of selected primers

S.NO	Primer	Sequence 5'-3'	Per cent polymorphism
1	UBC 811	GAGAGAGAGAGAGAGC	87.5%
2	UBC 841	GAGAGAGAGAGAGAYC	100%
3	UBC 857	ACACACACACACACACG	100%
4	UBC 826	ACACACACACACACACC	100%
5	UBC 815	CTCTCTCTTCTCTCTG	100%
6	UBC 888	BDBCACACACACACACA	100%
7	UBC 889	DBDACACACACACACAC	66.6%
8	UBC 891	HVHTGTGTGTGTGTGTG	100%
9	UBC 813	CTCTCTCTCTCTCTCTT	77.7%
10	ISSR M5	(GA)9-C	75%
11	ISSR-03	(CT)8-GC	100%
12	ISSR-04	(CA)6-AC	100%
13	ISSR-06	(CT)6-AG	100%
14	ISLA(CT)8TG	(CT)8-TG	83.3%
15	ISSR F2	(AG)8-CG	100%
16	UBC-809	AGAGAGAGAGAGAGAG	66.6%
17	UBC-810	GAGAGAGAGAGAGAGT	100%
18	UBC-823	TCTCTCTCTCTCTCTCC	50%
19	UBC-840	GAGAGAGAGAGAGAGT	62.5%
20	UBC-845	CTCTCTCTCTCTCTCTRG	100%
21	UBC-847	CACACACACACACACC	100%
22	UBC-864	ATGATGATGATGATGTG	87.5%
23	UBC-880	GGAGAGGAGAGGAGA	100%

Table 4
Details of unique bands displayed by different primers

SL.NO	PRIMER	GENOTYPE	UNIQUE BAND (base pairs)
1	ISSR-03	TSMLGB18	1000
2	ISSR-04	TSMLGB19	2500
3	UBC-810	TSMLGB18	800,750,600
4	UBC-823	TSMLGB18	1200,800
5	ISLA(CT) ₈ TG	TSMLGB19	2100
6	UBC-864	TSMLGB18	1200,900,800,500
7	ISSR-06	TSMLGB24	1300
8	UBC-811	TSMLGB17	1100,800
9	UBC-841	TSMLGB18	1400
10	UBC-889	TSMLGB1	1500
11	UBC-826	TSMLGB18	400
12	UBC-880	TSMLGB18	1300,1250,1000,900
13	ISSR M5	TSMLGB17	650,300
14	UBC-813	TSMLGB18	1500,900
15	UBC-891	TSMLGB18	2000,1000,100
16	UBC-840	TSMLGB24	300
17	UBC-809	TSMLGB18	1250,800,550,400
18	UBC-815	TSMLGB1	250
19	UBC-847	TSMLGB2	1100
		TSMLGB17	350
		TSMLGB18	2300,1600,1250,180
		TSMLGB6	1400
		TSMLGB18	350
		TSMLGB16	1400,900
		TSMLGB18	1750,500
		TSMLGB18	1500,1400,1000,750
		TSMLGB2	800
		TSMLGB13	1750

SL.NO	PRIMER	GENOTYPE	UNIQUE BAND (base pairs)
		TSMLGB2	650
		TSMLGB18	1200
		TSMLGB8	3100,500
		TSMLGB18	2000,500,300

Table 5
List of ISSR primers and their molecular characters

Primer	H	PIC	EMR	MI	D
UBC 811	0.46	0.51	1	0.46	0.86
UBC 841	0.48	0.51	1	0.48	0.67
UBC 857	0.49	0.50	1	0.49	0.70
UBC 826	0.42	0.54	1	0.42	0.51
UBC 815	0.42	0.53	1	0.42	0.52
UBC 888	0.45	0.52	1	0.44	0.56
UBC 889	0.32	0.57	1	0.31	0.35
UBC 891	0.34	0.57	1	0.34	0.38
UBC 813	0.38	0.55	1	0.37	0.43
ISSR M5	0.29	0.58	1	0.29	0.32
ISSR-03	0.34	0.56	1	0.34	0.39
ISSR-04	0.33	0.57	1	0.33	0.38
ISSR-06	0.24	0.59	1	0.24	0.26
ISLA(CT)8TG	0.34	0.57	1	0.34	0.34
ISSR F2	0.37	0.58	1	0.37	0.38
UBC-809	0.30	0.60	1	0.30	0.31
UBC-810	0.53	0.55	1	0.53	0.53
UBC-823	0.53	0.57	1	0.53	0.54
UBC-840	0.56	0.56	1	0.56	0.57
UBC-845	0.60	0.55	1	0.60	0.61
UBC-847	0.60	0.55	1	0.60	0.62
UBC-864	0.66	0.55	1	0.66	0.66
UBC-880	0.60	0.54	1	0.60	0.62

The Jaccard's co-efficient of similarity varied from 0.57(between TSMLGB18 and TSMLGB1) to 0.88 (between TSMLGB10 and TSMLGB11) (Fig. 5). A dendrogram based UPGMA analysis with ISSR data was shown in Fig. 5. The dendrogram divided the 24 genotypes into 7 clusters based on Jaccard's coefficient similarity matrix at 0.75 similarity. Cluster I consisted of TSMLGB1, TSMLGB13 and cluster II includes maximum number (15) of genotypes viz., TSMLGB3, TSMLGB4, TSMLGB6, TSMLGB8, TSMLGB5,

TSMLGB7, TSMLGB9, TSMLGB10, TSMLGB11, TSMLGB12, TSMLGB14, TSMLGB15, TSMLGB17, TSMLGB19 and TSMLGB20. Cluster III includes TSMLGB16, cluster IV has TSMLGB21, TSMLGB22, TSMLGB24. Cluster V, VI, VII consisted of single genotype each namely TSMLGB23, TSMLGB2, TSMLGB18 respectively.

Discussion

In the present investigation, ISSR markers were successfully employed in molecular diversity analysis to assess the phylogenetic relationships at the intraspecific level amongst 24 genotypes of *B. orellana* collected from the germplasm of TNAU, Mettupalayam, India. The 23 ISSR primers generated 249 bands among 24 genotypes, the molecular band size covered from 100 bp to 3500bp. The amplified fragment ranged from 6 (ISSR-O4) to 13 (UBC 826) with an average of 10.8 bands per primer. Similar results were reported in *Elaeocarpus serratus* (Raji R *et al.*, 2021), *Ricinus communis* (Gajera *et al.*, 2010), *Pistacia khinjuk* (Karimi *et al.*, 2012) and *Citrullus colocynthis* (Verma *et al.*, 2017) for the number of bands in ISSR markers. The primers have recorded high to moderate polymorphism indicating wider genetic diversity in the species. Similar results for ISSR markers have been reported in various woody species viz., *Jatropha curcas* (Singh *et al.* 2010), *Pinus roxburghii* (Rawat, 2014), *Pongamia pinnata* (Pradeep *et al.*, 2010), *Simarouba glauca* (Sekar *et al.*, 2003), *Albizia lebbek* (Radhakrishnan *et al.*, 2001), *Tectona grandis* (Parthibhan 2001).

The presence of unique bands facilitates the discrimination of the unknown genotypes, indicating that TSMLGB18 can be reliably identified based on its unique bands. The size of the amplified products varied from 100bp-3000bp and the PIC ranged between 0.6 – 0.5 (Table 3), signifying ISSR as an informative molecular marker. The high heterozygosity, multiplex ratio and discriminating power of markers revealed by the study exhibited the proficiency of ISSR markers for genetic diversity studies of *Bixa orellana*. The results were in line with the other species for marker values viz., *Elaeocarpus serratus* (Raji R *et al.*, 2021), *Acacia senegal* (Josiah *et al.* 2008), and *Populus cathayana* (Z lu *et al.*, 2006).

The Jaccard's co-efficient of similarity varied from 0.57 (between TSMLGB18 and TSMLGB1) to 0.88 (between TSMLGB10 and TSMLGB11) which indicated moderate level of genetic variation among the genotypes in the study. The highest JC value between TSMLGB10 and TSMLGB11 indicate their genetic similarity, whereas low JC between TSMLGB1 and TSMLGB18 denotes significant genetic differences. UPGMA cluster placed TSMLGB18 distantly from all other genotypes in the study possibly due to its distinctiveness whereas TSMLGB10 and TSMLGB11 placed closely at 0.88 similarity could be rendered to its similarity in genetic structure. Generally, the individuals from different geographic origins have different genetic constitution due to its natural selection for different traits (Raji *et al.*, 2021) and fall in same cluster based on their origin. The present study is a novel attempt to evaluate genetic diversity using ISSR markers in *Bixa orellana*.

The molecular analysis grouped the genotypes into 7 clusters. The wide divergence and a broad genetic base allow for identification and selection of divergent parents for any hybridization or improvement

plans and contribute significantly to the future breeding and conservation initiatives in the species. The study also revealed considerable variation in banding patterns, high heterozygosity and wide range of polymorphism among the genotypes, which demonstrated the suitability of ISSR markers as powerful tools for genetic characterization, diversity assessment and population structure analysis of *Bixa orellana* genetic resources. Previous studies on genetic diversity analysis using ISSR markers were reported on many species with similar results including *Santalum album* (Sandeep et al., 2020), *Eucalyptus cladocalyx* (Ballesta et al., 2015), *Elaeocarpus serratus* (Raji et al. 2021), *Morus* spp (Tikadar et al., 2003), *Tectona grandis* (Ansari et al., 2012), Bamboo spp (Amom et al., 2020) elucidating the effectiveness of ISSR markers in analysing the diversity.

Conclusion

This study presents a novel application of ISSR markers for assessing genetic diversity in *Bixa orellana*. The analysis revealed high levels of polymorphism, demonstrating the reliability and informativeness of ISSR markers for genetic characterization in this species. The substantial genetic variability observed highlights the heterogeneous nature of *B. orellana*, indicating considerable potential for heterosis in future breeding efforts. A well-designed breeding strategy could facilitate the development of superior progenies with enhanced bixin content and associated agronomic traits. These findings provide a valuable genetic framework for the selection of diverse and promising genotypes for future genetic advancement programs targeting bixin yield improvement.

Abbreviations

Declarations

Acknowledgment;

The infrastructure support provided by Forest College and Research Institute, Mulugu, is greatly appreciated by the authors. We are very grateful to FC&RI, Mettupalayam, Tamilnadu for supplying the seed material on schedule.

Conflict of interest: Authors do not have any conflict of interests to declare.

Ethical issues: None.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contribution

ISSR	:	Inter Simple Sequence Repeats
SSR	:	Simple Sequence Repeats
UPGMA	:	Unweighted pair group mean arthematics
DNA	:	De-oxy ribo Nucleic Acid
CTAB	:	Cetyl trimethyl ammonium bromide
UBC	:	University of British Columbia
PCR	:	Polymerase Chain Reaction
ISSR	:	Inter simple sequence repeats
UV	:	Ultraviolet
SIMQUAL	:	Similarity for Qualitative data
SCAR	:	Sequence Characterized Amplified Region
TNAU	:	Tamilnadu Agriculture University
TBE	:	Tris/Borate/EDTA
kb	:	Kilobytes
ml	:	Milli-litres
µl	:	Micro-litres
nm	:	Nano-metres
min	:	Minutes
%	:	Percentage
°C	:	Degree-centigrade

Manasa conducted the entire research as part of her postgraduate studies. Reeja framed the research objectives, design and methodology and provided guidance. Lalitha supported the research execution, in field and lab, while Arjun contributed to the selection of primers and supported statistical interpretations. Shalini participated in design and methodology. Kumaran provided the seed material from germplasm bank, Tamil Nadu. Sreedhar assisted in statistical analysis and Chiranjeeva provided technical guidance during the seedling nursery stage. All authors have read and approved the final manuscript.

References

1. Alcázar-Alay S. C., Osorio-Tobón J. F., Forster-Carneiro T., and Meireles M. A. A. (2017) Obtaining bixin from semi-defatted annatto seeds by a mechanical method and solvent extraction: process integration and economic evaluation. *Food research international*,99: 393–402.
2. Amiryousefi A, Hyvönen J, Poczai P. (2018) iMEC: Online Marker Efficiency Calculator. *Appl Plant Sci.* Jun 24;6(6):e01159. doi: 10.1002/aps3.1159. PMID: 30131901; PMCID: PMC6025818.
3. Amom T., Tikendra L., Apana N., Goutam M., Sonia P., Koijam A.S. and Nongdam P. (2020) Efficiency of RAPD, ISSR, iPBS, SCoT and phytochemical markers in the genetic relationship study of five native and economical important bamboos of North-East India. *Phytochemistry*, 174: 112–330.
4. Ansari S. A., Narayanan C., Wali S. A., Kumar R., Shukla N. and Rahangdale S. K. (2012) ISSR markers for analysis of molecular diversity and genetic structure of Indian teak (*Tectona grandis* Lf) populations. *Annals of Forest Research*,55(1): 11–23.
5. Ballesta P., Mora F., Contreras-Soto R. I., Ruiz E. and Perret S. (2015) Analysis of the genetic diversity of *Eucalyptus cladocalyx* (sugar gum) using ISSR markers. *Acta Scientiarum. Agronomy*, 37:133–140.

6. Botstein, D., White R. L., Skolnick M., and Davis R. W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* 32: 314–331.
7. Dequigiovani G., Ramos S. L. F., Alves-Pereira A., Fabri E. G., Carvalho P. R. N., da Silva M. G. and Veasey E. A. (2017) Genetic diversity and structure in a major Brazilian annatto (*Bixa orellana*) germplasm bank revealed by microsatellites and phytochemical compounds. *Genetic Resources and Crop Evolution*, 64(7): 1775–1788.
8. Doyle, J. (1991) DNA protocols for plants. In *Molecular techniques in taxonomy* (pp. 283–293). Berlin, Heidelberg: Springer Berlin Heidelberg.
9. Franco C.F.O., Fabri E.G., Barreiro Neto M. (2008). Urucum: sistemas de produção para o Brasil. EMEPA, João Pessoa
10. Gajera BB, Kumar N, Singh AS et al (2010) Assessment of genetic diversity in castor (*Ricinus communis* L.) using RAPD and ISSR markers. *Ind Crops Prod* 32(3):491–498.
11. <https://irscope.shinyapps.io/iMEC/>
12. Josiah C. C., George D. O., Eleazar O. M. and Nyamu W. F. (2008) Genetic diversity in Kenyan populations of *Acacia senegal* (L.) willd revealed by combined RAPD and ISSR markers. *African Journal of Biotechnology*, 7(14).
13. Karimi HR, Abadi MHH, Kohbanani AM (2012) Genetic diversity of *Pistacia khinjuk* Stocks. using RAPD markers and leaf morphological characters. *Plant Syst Evol* 298(5):963–968
14. Kumaran, K. (2014) Production potential of annatto (*Bixa orellana* L.) as a source of natural edible dye. In *the proceedings of the International workshop on natural dyes, Hyderabad, India*.
15. Leal F., de Clavijo CM.(2012).Annatto: botany and horticulture. In: Janick J (ed) *Horticultural Reviews*. Wiley, Hoboken. 389–419.
16. Li G., and Quiros C. F. (2001) Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in Brassica. *Theoretical and applied genetics*, 103(2), 455–461.
17. Liu, B. H. (1998) *Statistical genomics: Linkage, mapping and QTL analysis*. CRC Press, Boca Raton, Florida, USA.
18. Lu Z., Wang Y., Peng Y., Korpelainen H. and Li C. (2006) Genetic diversity of *Populus cathayana* Rehd populations in southwestern China revealed by ISSR markers. *Plant Science*, 170(2): 407–412.
19. Parthiban K. T. (2001) Seed source variations, molecular characterization and clonal propagation in teak (*Tectona grandis* Linn. F.). *PhD, Tamil Nadu Agricultural University, Coimbatore, India*.
20. Pattanaik L., Naik S. N. and Hariprasad P. (2019) Valorization of waste *Indigofera tinctoria* L. biomass generated from indigo dye extraction process—potential towards biofuels and compost. *Biomass Conversion and Biorefinery*, 9(2): 445–457.
21. Pradeep kumar T., Hegde V., Sujatha R. and George T.(2010)Assessment of *Pongamia pinnata* (L.)–a biodiesel producing tree species using ISSR markers. *Current Science*, 99(10):13–27.

22. Radhakrishnan S. (2001) Genetic divergence and DNA based molecular characterization in *Albizia lebbeck* (L.) Benth (Doctoral dissertation, Ph. D. (For.) Thesis, Tamil Nadu Agricultural University, Coimbatore).
23. Raji R. and Siril E. A. (2021) Genetic diversity analysis of promising Ceylon olive (*Elaeocarpus serratus* L.) genotypes using morphological traits and ISSR markers. *Current Plant Biology*, 26: 100–201.
24. Rawat A., Barthwal S., and Ginwal, H. S. (2014) Comparative assessment of SSR, ISSR and AFLP markers for characterization of selected genotypes of Himalayan Chir pine (*Pinus roxburghii* Sarg.) based on resin yield. *Silvae Genetica*, 63;1–6:94–108.
25. Sandeep C., Kumar A., Rodrigues V., Viswanath S., Shukla A. K. and Sundaresan V. (2020) Morpho-genetic divergence and population structure in Indian *Santalum album* L. *Trees*, 34(5): 1113–1129.
26. Sekar I. (2003) *Seed storage, seed source variations, molecular characterization and in vitro propagation in Simarouba glauca DC* (Doctoral dissertation, Ph. D. Thesis, Tamil Nadu Agricultural University, Coimbatore, India).
27. Singh P., Singh S., Mishra S. P. and Bhatia S. K. (2010) Molecular characterization of genetic diversity in *Jatropha curcas* L. *Genes Genomes Genomics*, 4:1–8.
28. Smith J., and Wallin H. (2006) Annatto extracts. Chemical and technical assessment
29. Tessier, C., David J., This P., Boursiquot J. M., and Charrier A. (1999) Optimization of the choice of molecular markers for varietal identification in *Vitis vinifera* L. *Theoretical and Applied Genetics* 98: 171–177.
30. Tikader A., Rao A. A. and Thangavelu K. (2003) Genetic divergence in exotic mulberry (*Morus* spp.) germplasm. *Sericologia*, 43(4): 495–501.
31. Verma, K. S., ul Haq, S., Kachhwaha, S., & Kothari, S. L. (2017) RAPD and ISSR marker assessment of genetic diversity in *Citrullus colocynthis* (L.) Schrad: a unique source of germplasm highly adapted to drought and high-temperature stress. *3 Biotech*, 7, 1–24.
32. Vilares A.S., Sa~o Jose´ A.R., Rebouc T.N.H., Souza I.V.B. (1992) Estudo da biologia floral de urucuzeiro (*Bixa orellana* L.). *Rev Bras Corantes Nat*, 1:101–105
33. Vos P., Hogers R., Bleeker M., Reijans M., Lee T. V. D., Hornes M., and Zabeau M. (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic acids research*, 23(21):4407–4414.
34. Williams J.G., Kubelik A.R., Livak K.J., Rafalski J.A., Tingey S.V. (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic acids research*, 18(22): 6531–6535.
35. Zhang, Y., Zhang, X., Wang, Y. H., & Shen, S. K. (2017) De novo assembly of transcriptome and development of novel EST-SSR markers in *Rhododendron rex* Lévl. through Illumina sequencing. *Frontiers in Plant Science*, 8, 1664.

Figures

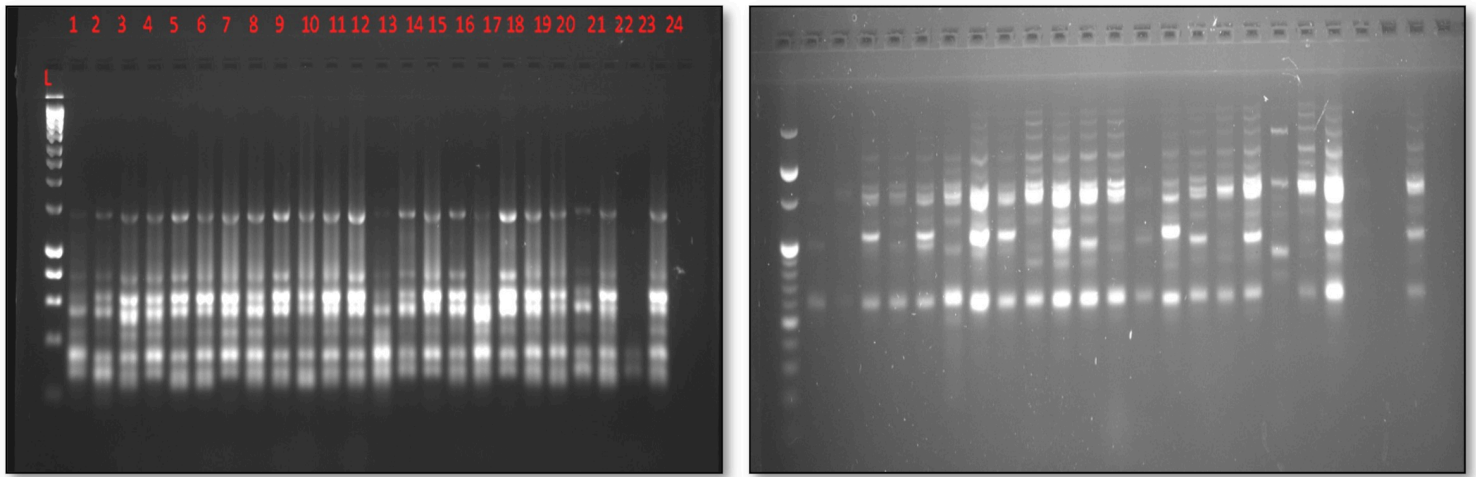


Figure 1

ISSR profile deciphering 100% polymorphism for primers a) ISSR F2 and b) UBC 845; Lane L is the 1kb DNA ladder; Lane 1 to 24 indicates the DNA of genotypes under study.

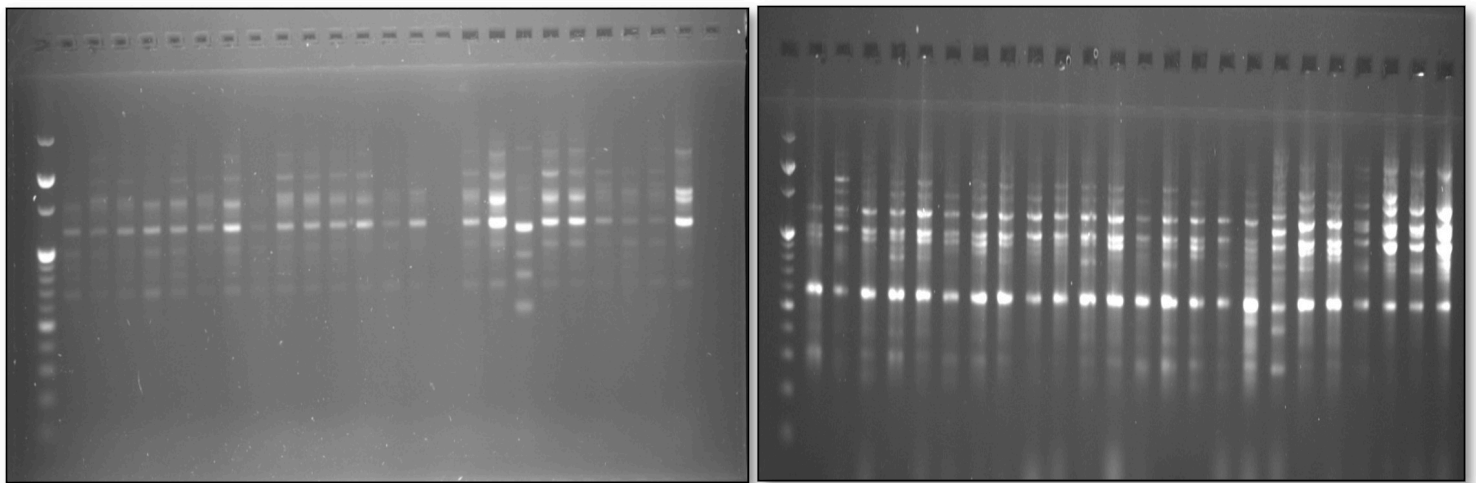


Figure 2

ISSR profile deciphering minimum polymorphism for primers a) UBC 823 and b) UBC 864 with 50%&62.5% respectively; Lane L is the 1kb DNA ladder; Lane 1 to 24 displayed the DNA of genotypes under study.

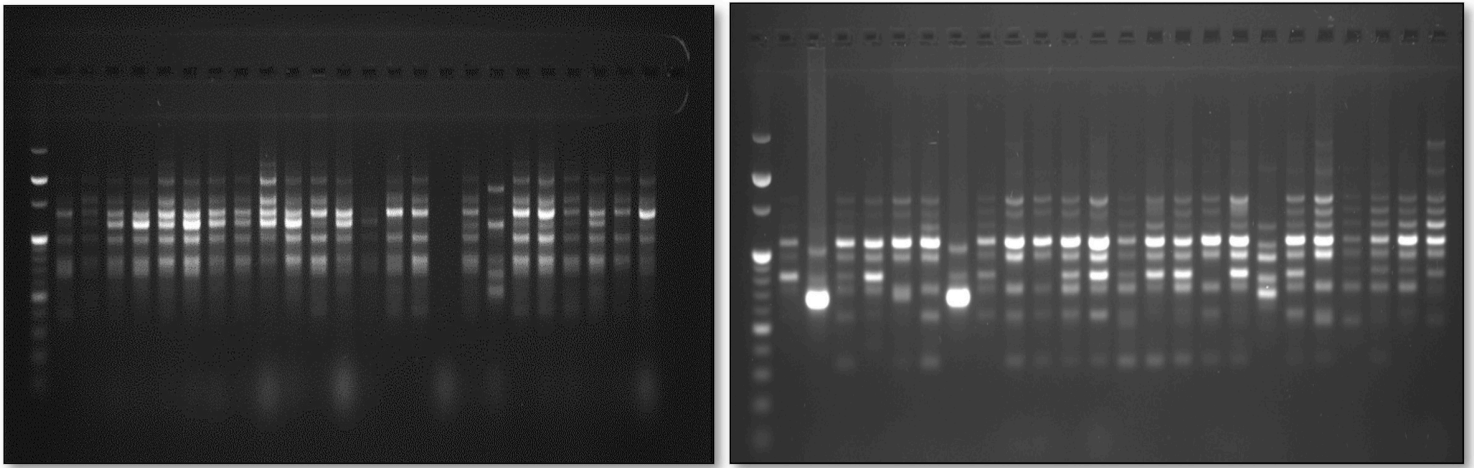


Figure 3

ISSR profile produced with unique bands in TSMLGB18 for a)1750bps for UBC 813 and b) ISSR-O4 at 800 bps; where lane L is 1kb ladder; numbers from 1 to 24 are genotypes DNA under study

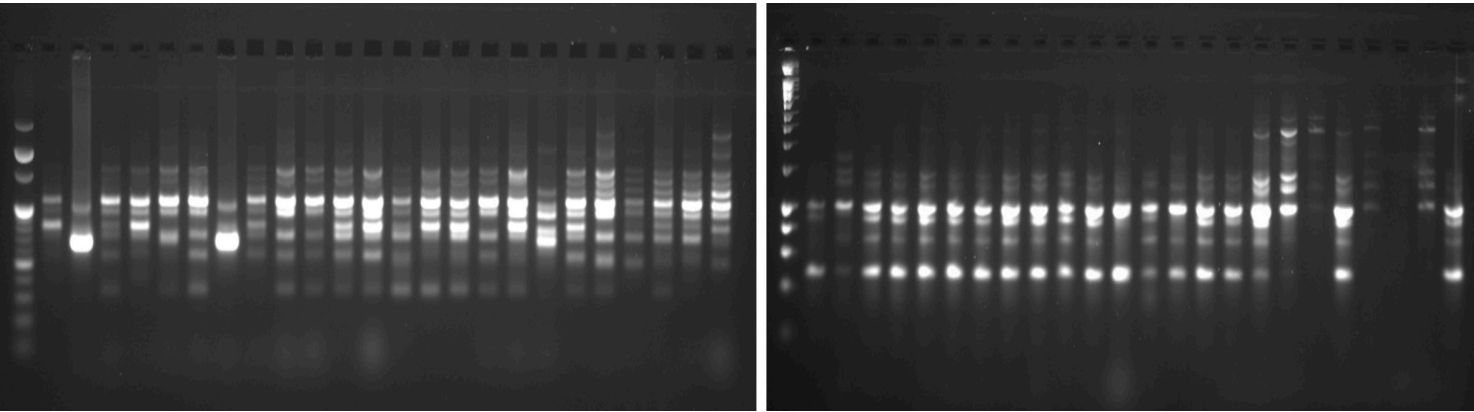


Figure 4

ISSR profile produced with unique bands in TSMLGB2 for a) 1100 bps for UBC 826 and b) UBC 840 at 800 bps; where lane L is 1kb ladder; numbers from 1 to 24 are genotypes DNA under study

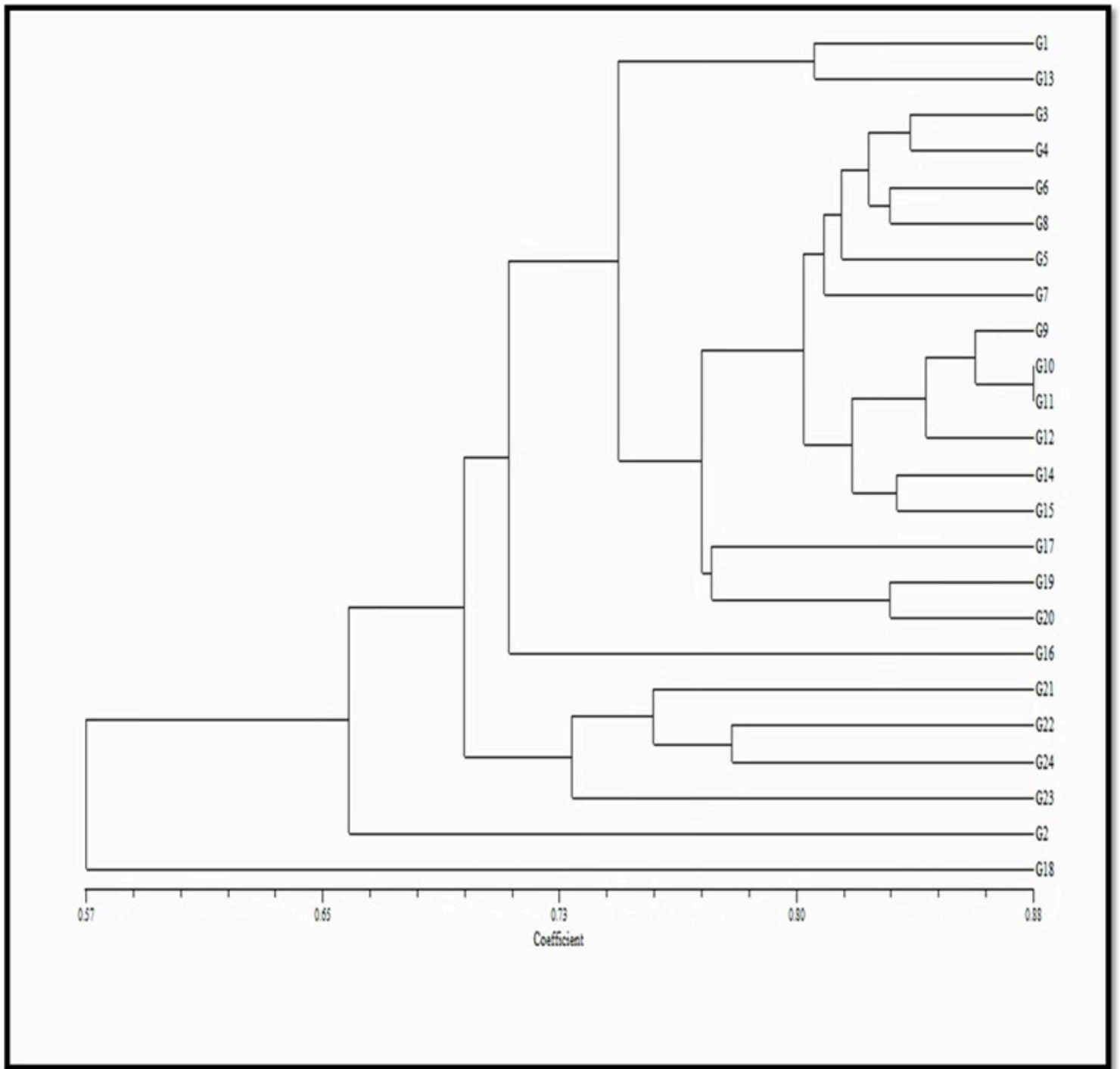


Figure 5

UPGMA dendrogram computed using genetic distance matrix based on ISSR data