Plant Physiol. Rep. (July–September 2022) 27(3):407–418
https://doi.org/10.1007/s40502-022-00675-8

ORIGINAL ARTICLE

Understanding the genetic determinants and population structure of *Pongamia pinnata* (L.) Pierre for oil yield and its properties using transcriptome derived SSR markers

K. Rajarajan[1] · S Sakshi[1] · M. Rana[2] · A. Radhakrishna[2] · R. Vishnu[1] · H. Anuragi[1] · A. K. Handa[1] · A. Arunachalam[1]

Received: 11 January 2022 / Accepted: 11 July 2022 / Published online: 6 August 2022

Abstract *Pongamia* is a commercially important tree; its seed oil has the potential to be used as a biofuel worldwide. The seed traits have huge genetic diversity, which suggests its scope for genetic improvement. However, lack of improved genetic stock is a major bottleneck to establishing plantations for biodiesel production. Assessment of genetic diversity, population genetic structure, and biodiesel properties in 18 *Pongamia* accessions were studied using eight morphological traits and 11 SSR markers. The accessions exhibited a high degree of genetic variability at both levels, indicating the potential scope of this population for further tree breeding. Certain morphological traits are governed by additive genes and may provide indirect selection to yield improvement. Based on the dissimilarity of the coefficient, NRCP10 and NRCP20 were identified as the most distinct accessions, and, interestingly, these two accessions also have a higher cetane index for better biofuel properties. The dendrogram divided 18 accessions into two major clusters. The principal component analysis revealed a total of 70.92% variability in this population. On the other hand, molecular characterization using eleven SSR markers exhibited a significant level of polymorphism across the accessions. A dendrogram based on SSR markers clearly separated the 18 accessions into five groups, which resulted in NRCP11, NRCP14 and NRCP21 being the most distinct accessions. Also, the dendrogram conforms to the first and second coordinates by principal component analysis. The population structure of these 18 accessions also revealed the presence of two gene pools. In addition, analysis of molecular variance found 97% of the variability within the population, while only 3% occurred between populations. The genotypes identified in this study using morphological and molecular markers could be selected as donor parents in *Pongamia* tree breeding programs to improve economic trait values.

Keywords *Pongamia pinnata* · Genetic diversity · Population structure · SSR markers

✉ K. Rajarajan
krajarajantnau@gmail.com

S Sakshi
sakshisahu37@gmail.coms

M. Rana
maneet.rana@icar.gov.issn

A. Radhakrishna
AR.Krishna@icar.gov.in

R. Vishnu
VISHNU.R@icar.gov.in

H. Anuragi
hirdayesh.anuragi@icar.gov.in

A. K. Handa
Arun.Handa@icar.gov.in

A. Arunachalam
a.arunachalam@icar.gov.in

[1] ICAR-Central Agroforestry Research Institute, Jhansi, India

[2] ICAR-Indian Grassland and Fodder Research Institute, Jhansi, India

Introduction

Pongamia pinnata (L.) Pierre is a diploid (2n = 22) multipurpose tree species (Fabaceae) native to India and Southeast Asia (Al Muqarrabun et al., 2013; Karmee & Chadha, 2005; Kesari et al., 2009). Subsequently, it was introduced in to China, New Zealand, Australia, and the United States (Scott et al., 2008). In agroforestry programs, this species has been widely considered due to its ease of

propagation in marginal lands, nitrogen-fixing capacity, used as green manure, fast-growth attributes, and, most importantly, high seed oil yield (Pavithra et al., 2014; Scott et al., 2008). This tree has also gained momentum across the world due to its biodiesel properties (Karmee & Chadha, 2005). Additionally, it has several important applications in medicine as an antimicrobial and in agriculture as a bio-insecticide and nematicide (Shivanna & Rajakumar, 2010; Yadav et al., 2004).

Increasing fuel product prices will probably have negative consequences for the global automobile industry in the future, which drives a search for alternative renewable energy sources. In this search, *Pongamia* is one of the promising sources of non-edible oil yielding trees that could be considered for biodiesel production (Azam et al., 2005; Karmee & Chadha, 2005; Kesari & Rangan, 2010). The seed oil content (30–40%) is one of the important traits responsible for its commercial usage, since it can be translated into biodiesel by trans-esterification (Karmee & Chadha, 2005; Naik et al., 2008). The major impediment to efficiently exploiting the biodiesel potential of this species is a lack of improved genetic stocks (Sharma et al., 2016). Genetic diversity estimation is an imperative step for successful tree breeding (Kesari et al., 2009). It allows tree breeders to have efficient parental selection for hybridization programs (Badu-Apraku et al., 2021). Phenotypic characterization of bigger populations allows germplasm assemblage and grouping, which aids in the selection of promising genotypes for tree breeding. Despite its importance, it has certain constraints in plant breeding since it is strongly influenced by the environment and is labor intensive. Considering morphological traits has limited value, DNA markers may be a better choice for genotyping (Uchoi et al., 2016). Likewise, analysing a population structure is more important in the context of genetic diversity and association mapping studies (Kumar et al., 2020). It also provides information on associations between markers and traits of interest. In this study, the population structure was assessed using SSR markers. Unlike other molecular markers, SSR has a high rate of polymorphism, co-dominant, multi-allelic, and wide genome coverage (Adjebeng-Danquah et al., 2020; Rajarajan & Ganesamurthy, 2011; Rajarajan et al., 2021; Rohini et al., 2020; Uchoi et al., 2016). Despite the advancement of molecular markers, there have been few studies on the genetic variation in *Pongamia* genotypes for seed and oil yield traits (Kesari & Rangan, 2011; Kesari et al., 2010; Sharma et al., 2016). There has been no significant research combining morphological and SSR markers to understand the genetic diversity, population structure, and biodiesel properties of *Pongamia*. However, characterizing genetic diversity of *Pongamia pinnata* gained significant concern due to socioeconomic prominence.

Therefore, this study aimed to investigate the genetic diversity, population structure, and biodiesel properties in *P. pinnata* using morphological and molecular markers for its genetic improvement.

Materials and methods

Plant material

Eighteen candidate plus trees (accessions) of *Pongamia pinnata* were collected from four states in India viz. Madhya Pradesh, Uttar Pradesh, Haryana, and Rajasthan (Table 1). The progeny trial was established in 2005 at Central Agroforestry Research Institute, Jhansi (Ahlawat et al., 2016). The plus trees were selected on the basis of high fruiting, seed yield, clean bole, and close canopy. These progenies were raised in a randomized block design (RBD) for three replications with five plants of each progeny per replication. Spacing of 5 m × 5 m was maintained in the field.

Oil and biodiesel properties

The seeds of each plus tree were combined to represent an entry, which was then analysed using a Soxhlet apparatus (Ambassador (BPI-32) Soxhlet Extraction Unit, New Delhi) using petroleum ether as the solvent. The 3 g of

Table 1 Details of *Pongamia pinnata* accessions used in this study

No	Accession name	Collection site	State
1	NRCP6	Basi, Laltipur	Uttar Pradesh
2	NRCP7	Basi, Laltipur	Uttar Pradesh
3	NRCP9	Babina, Jhansi	Uttar Pradesh
4	NRCP10	Matatila Dam, Lalitpur	Uttar Pradesh
5	NRCP11	Matatila Dam, Lalitpur	Uttar Pradesh
6	NRCP12	Malthola, Lalitpur	Uttar Pradesh
7	NRCP13	Barodia, Laltipur	Uttar Pradesh
8	NRCP14	Nandan bara, Laltipur	Uttar Pradesh
9	NRCP16	Bhagirathpur, Dholpur	Rajasthan
10	NRCP17	Tekanpur, Gwalior	Madhya Pradesh
11	NRCP18	Tekanpur, Gwalior	Madhya Pradesh
12	NRCP20	Bassi, Jaipur	Rajasthan
13	NRCP21	Chhotowaara, Bharatpur	Rajasthan
14	NRCP22	Panipat	Haryana
15	NRCP23	Shalepur, Dholpur	Rajasthan
16	NRCP24	Ramnath City, Jhansi	Uttar Pradesh
17	NRCP25	Matatila, Jhansi	Uttar Pradesh
18	NRCP26	Mauranipur, Jhansi	Uttar Pradesh

seed powder and 150 mL of petroleum ether were used for each set. Extraction was done at 80 °C for 5 h, followed by a rapid solvent recovery at 120 °C for 60 min. The mean of three replicates of each accession was calculated. Each accession's oil content was determined and given as a percentage (w/w).

Saponification number (SN) and iodine value (IV) were estimated using FAMEs composition of oil as described by Kalayasiri et al., (1996).

$$SN = \sum \left(560 \times A_i\right)/MW_I$$

where A_i is the percentage, D is the number of double bonds MW_i is the molecular weight of each component.

$$IV = \sum \left(254 \times D \times A_i\right)/MW_i$$

According to Krisnangkura (1986), the Cetane Number (CN) of FAMEs was computed for estimating the biodiesel properties of each individual tree accession's oil.

$$CN = 46.3 + 5458/SN - 0.225 \times IV$$

Morphological characterization

The morphological traits, viz. pod, seed, and oil characters, were studied in the present study for 18 accessions. Each morphological traits were calculated from three replications.

Pod traits

The pod length (PDL), breadth (PDB) and thickness (PDT) were calculated using calipers. The length of the pod was measured from the tip to the base; the width of the pod was measured from left to right; the thickness of the pod was determined on the part that contains seeds. The pot traits are expressed in mm.

Seed and oil traits

The seed traits such as seed length (SDL), breadth (SDB) and thickness (SDT) was calculated using caliper. The length of the seed was measured from the tip to the base; the width of the seed was measured from left to right; the thickness of the seed was observed from the front to the back and expressed in mm. While, 100-seed weight was measured by counted 100 seeds and weighted and expressed in grams. The oil content for each accession was calculated and expressed as percentage (w/w) of dry seed using following equation

$$\text{Percent oil content} = \left[(Wfb - Wib)/Ws\right] \times 100$$

where, Wfb is final beaker weight, Wib is initial beaker weight and Ws is the sample weight.

Statistical analysis

Simple statistical measures, such as mean and standard error, were considered in the quantitative data analyses. INDOSTAT (Indostat Services, Hyderabad, India) was used to analyze genetic variability measures: genotypic coefficient of variation (GCV), phenotypic coefficient of variation (PCV), heritability (h^2) and genetic advance as percent mean (GAM). The dendrogram was constructed based on the group average clustering method, and principal component analysis was performed to find relationships among the variables and accessions using XLSTAT (Addinsoft, Paris), based on the morphological attributes.

Molecular marker analysis

The genomic DNA was isolated from young leaves using the modified CTAB (Cetyl trimethyl ammonium bromide) method (Doyle & Doyle, 1990). DNA purity was assessed by UV-Spectrophotometer (Eppendorf, Germany). A total of 20 SSR repeats were randomly selected from the assembled Pongamia transcriptome sequences in NCBI (SRA046342.1) (Huang et al., 2012), using MIcroSAtellite (MISA, http://pgrc.ipk-gatersleben.de/misa). The minimum number of repeats used to select the SSRs was 7 for di-nucleotide repeats, 4 for tri-nucleotide repeats, and 3 for tetra, penta and hexanucleotide repeats. Primer pairs for each SSR were designed using the standalone Primer3 (http://primer3.sourceforge.net/). Out of these markers, only 11 primers were found to be polymorphic (Table 2). Amplification was carried out in 20 µl volumes containing 2 µl genomic DNA (20 ng), 2 µl of 10 × Taq Buffer with $MgCl_2$, 1 µl of dNTP (10 mM), Taq DNA polymerase (3.0 U/µl), 1 µl of each forward and reverse primer (0.2 mM), and 12.7 µl of nuclease-free water. The thermal cycler (Eppendorf Master-Cycler, Germany) with thermal cycling program follows: initial denaturation for 3 min at 93 °C, followed by 35 cycles of denaturation for 30 s at 93 °C, with an optimized annealing temperature for 30 s (54.6–56.3 °C), then extension for one minute at 72 °C, and the final extension for 8 min at 72 °C. The amplified fragments were separated on 3% agarose gel using 80 V for 2 h by gel electrophoresis (Syngene, USA).

Data analysis

The amplified fragments were used to construct a binary matrix (presence/absence of bands), to generate a dendrogram based on the unweighted neighbor-joining method, and principal coordinate analysis using DARwin software (Perrier & Flori, 2003). According to Botstein et al., (1980), each primer set was calculated for polymorphic

Table 2 List of SSR markers used in this study

Primer code	Repeat motif	Forward sequence	Reverse sequence	Annealing temperature (°C)
P001	(CT)17(CTAT)4	GAAGAGTAACCATGGTGGCAA	TCATACACCACCAAACGAGC	55.3
P002	(TCTTT)3(TTTC)3tt(AGCTTC)3	CGTCACGTTTCCTTAGCTCC	TCGAATCCTTCCATCCAGTC	54.9
P003	(GAG)7(GAGAAG)3	CGGACCCAATCTCCATTAGA	GGCAGAGGGAAGGGAATAAG	56.0
P004	(CA)12	ACCTTGACTCCATCCACAGC	CTTCTTGCATCAGGGTAGGC	55.8
P005	(AC)11	AACACTTGAGGCATTGGACC	CTTTAGCACTTGGCCTCTGG	55.0
P006	(AT)11	AGCTAGGTGCAGTCCTTCCA	AGACACAGCAATCACATGCC	54.3
P007	(GCG)8	AAAAGGACAAGAGGCGTTCA	GGCGCTCTTTCTCTCTCTCA	56.3
P008	(GGA)8	AGACGGTGGAGATGCTATGG	AGGTCTCGCTTGGAAACTCA	55.9
P009	(TCG)8	CGGCTTTGTTCCCATCTTTA	CGACTCCAAGGAGGTCAGAG	55.3
P010	(ATTG)6	ACAATGGTGCCCAGAAGAAC	GCACTGGGGTTTCTTGGTAA	55.2
P011	(TAGT)6	ACTCATCCGAAAACCGTCAG	ATTGCAAGCACCACAGTTCA	55.3

information content (PIC). STRUCTURE software (v.2.3.3) was used to study the genetic structure using a Bayesian model (Falush et al., 2007). Also, STRUCTURE HARVESTER (Irfan et al., 2010) was used to resolve the estimated value of Ln probability of data-LnP (K). K, which ranged from 0 to 10, was calculated using three iterations of the data in this analysis. GenAlex 6.4 was used to analyze molecular variance (AMOVA) according to Peakall and Smouse, (2012).

Results and discussion

Morphological characterization of pod, seed and oil traits

The mean values of the eight morphological traits in the 18 accessions exhibited significant variation (Table 3). In NRCP13, the maximum pod length and pod breadth were observed, while the minimum pod length was found

Table 3 Mean performance for pod, seed and oil traits among 18 *Pongamia* accessions

Accessions	PDL (mm)	PDB (mm)	PTC (mm)	SDL (mm)	SDB (mm)	STC (mm)	SDW (g)	OC (%)
NRCP6	53.037 ab	26.277 a	9.903 bc	21.960 bc	16.663 ab	7.367 b	118.077 b	37.260 ab
NRCP7	47.590 cde	23.130 cd	6.700 g	19.847 cde	15.070 de	6.133 fg	103.800 d	30.930 d
NRCP9	53.323 a	26.530 a	10.053 b	23.230 ab	16.673 ab	8.123 a	127.360 a	37.723 a
NRCP10	48.310 bcde	21.940 cd	9.747 bcd	20.057 cde	16.673 ab	6.430 def	65.600 g	33.787 bcd
NRCP11	48.940 abcde	26.530 a	8.277 ef	18.307 def	13.583 f	6.123 fg	49.000 h	33.640 bcd
NRCP12	45.187 e	17.570 e	8.897 de	18.227 ef	16.277 abcd	6.467 def	113.980 bc	32.370 cd
NRCP13	53.490 a	27.750 a	11.353 a	24.247 a	17.043 a	8.550 a	128.897 a	38.143 a
NRCP14	48.390 bcde	23.183 cd	8.950 cde	20.790 c	15.583 bcd	8.120 a	117.040 b	35.903 abc
NRCP16	50.093 abcd	17.587 e	8.100 ef	21.097 bc	15.370 cd	6.303 efg	69.880 fg	33.150 cd
NRCP17	36.760 f	26.800 a	7.870 f	19.707 cde	14.093 ef	6.307 efg	100.940 d	33.640 bcd
NRCP18	40.337 f	21.547 d	8.943 cde	20.810 c	15.737 bcd	5.810 g	76.357 f	34.570 abcd
NRCP20	50.573 abcd	23.980 bc	9.657 bcd	21.557 bc	16.337 abc	6.860 cd	114.160 bc	35.590 abc
NRCP21	39.960 f	17.267 e	6.693 g	20.540 c	15.207 cde	6.163 efg	105.680 cd	35.900 abc
NRCP22	51.987 abc	25.447 ab	9.743 bcd	21.870 bc	16.467 abc	7.300 bc	117.040 b	35.590 abc
NRCP23	40.193 f	22.733 cd	10.057 b	21.403 bc	13.123 f	5.977 fg	49.847 h	35.280 abc
NRCP24	46.340 de	21.657 cd	9.227 bcd	20.457 cd	16.660 ab	6.847 cd	48.950 h	32.980 cd
NRCP25	48.230 bcde	19.050 e	9.660 bcd	20.073 cde	16.290 abcd	6.687 de	87.837 e	34.637 abcd
NRCP26	46.450 de	18.290 e	7.590 f	17.213 f	13.590 f	6.247 efg	75.200 f	32.707 cd

Mean values of PDL- Pod length, PDB- Pod breadth, PTC- Pod thickness, SDL- Seed length, SDB- Seed breadth, STC- Seed thickness, SDW- 100 seed weight and OC- Oil content. Different letters indicates a significant differences ($p < 0.05$) by Duncan's test

in NRCP17. Similarly, the minimum pod breadth was recorded in NRCP21. The thickness of the pods also varied significantly across the 18 accessions. The maximum seed length, seed breadth, and 100-seed weight were observed in NRCP13. Seed thickness varied significantly across the accessions. Based on the mean performance for pod and seed traits NRCP13 has superior performed than other accessions. Similarly, Kaushik et al., (2007) has reported *Pongamia* CPT-33 had maximum values for most of the pod and seed traits. In addition, the oil content value in this study ranged from 30.93 percent (NRCP7) to 38.15% (NRCP13) in Table 3. The oil content showed high variation among the genotypes as it was evident from the range of mean values.

Biofuel properties

Across the 18 *Pongamia* accessions, there were substantial variations observed in the biofuel traits, such as saponification number (SN), iodine value (IV), and cetane index (CI) (Table 4). The maximum saponification number and Iodine value were found in NRCP11 and NRCP23, while the minimum SN and IV were found in NRCP20. Nevertheless, NRCP20 has the highest cetane index. In this study, the saponification number and the cetane index obtained are similar to the previous report (Sharma et al., 2016). According to the Indian biodiesel standard (IS15607: 2005), the minimum cetane number should be 51; however, in this study the average cetane number was 53.49 showing that these accessions exhibit high biodiesel potential (Mukta et al., 2009; Sharma et al., 2016).

Table 4 List of 18 *Pongamia* accessions mean values for biodiesel properties

Accessions	Saponification number	Iodine value	Cetane index
NRCP6	230.255 abc	89.110 ab	48.505 d
NRCP7	211.745 abcde	85.255 abc	54.310 bcd
NRCP9	199.270 de	82.700 abc	55.255 abc
NRCP10	212.625 abcde	81.865 abc	53.285 bcd
NRCP11	237.605 a	89.115 ab	50.725 bcd
NRCP12	207.545 bcde	85.465 abc	54.590 bcd
NRCP13	204.290 cde	81.445 bc	54.835 bcd
NRCP14	223.450 abcd	87.890 ab	52.850 bcd
NRCP16	214.950 abcde	77.930 cd	55.100 abc
NRCP17	199.250 de	78.315 cd	57.130 ab
NRCP18	210.685 abcde	78.895 cd	56.855 ab
NRCP20	193.510 e	71.435 d	61.160 a
NRCP21	234.290 ab	86.015 abc	51.480 bcd
NRCP22	218.195 abcde	79.695 c	53.040 bcd
NRCP23	237.510 a	89.880 a	48.705 cd
NRCP24	222.505 abcd	83.415 abc	51.245 bcd
NRCP25	230.925 abc	85.260 abc	52.175 bcd
NRCP26	234.290 ab	89.625 ab	51.580 bcd

Mean values for biodiesel properties. Different letters indicates a significant differences ($p < 0.05$) by Duncan's test

Table 5 Genetic variability for yield associated morphological traits of *Pongamia pinnata*

Traits	Coefficient of variation %		Heritability %	Genetic advance as percent of mean (%)
	Phenotypic	genotypic		
Pod length	10.70	10.36	93.72	20.66
Pod Breadth	15.02	14.02	87.07	26.94
Pod thickness	13.97	13.36	91.47	26.32
Seed length	8.70	7.89	82.27	14.74
Seed Breadth	7.93	7.50	89.31	14.59
Seed thickness	11.40	11.08	94.51	22.20
100 seed weight	30.50	29.95	96.45	60.60
Oil content	6.65	6.50	95.42	13.08

Morphological trait based genetic variability and diversity

In this study, the GCV was less than the PCV for pod, seed, and oil content traits indicating that non-additive gene action predominated (Table 5). Estimation of broad-sense heritability showed maximum heritability for the 100-seed weight followed by oil content. Table 5 also shows maximum genetic advance for the 100-seed weight followed by pod breadth. Based on these results, the 100-seed weight, seed thickness, pod length, pod breadth, and pod thickness had high heritability and genetic advance over other traits, indicating these traits are influenced by additive genes. In addition, oil yield had high heritability coupled with medium genetic advance indicating the additive gene effects on oil trait and as a result, the selection of these traits would be effective for oil and other yield associated trait improvement. Furthermore, the results in this study correspond with Ahlawat et al., (2016) and Kaushik et al., (2007) for *Pongamia* pod and seed traits.

The Euclidean distance-based cluster analysis used in this study grouped 18 accessions into two major clusters (Fig. 1a). There were ten accessions in Cluster I and eight in Cluster II. Cluster I was further subdivided into 1A and 1B, with six and four accessions, respectively. Cluster II was separated into groups 2A and 2B having three and five accessions, respectively. Furthermore, among the accessions, NRCP20 and NRCP10, which had the highest dissimilarity index, were shown to be the most diversified. Ahlawat et al., (2016) found a similar pattern of clustering

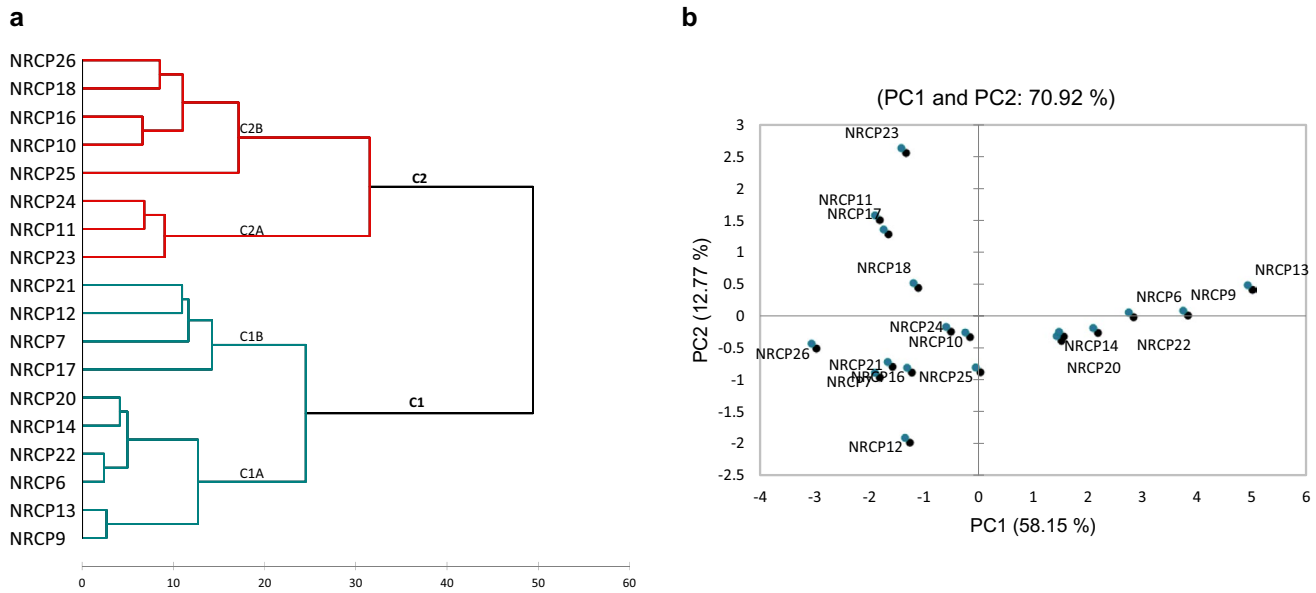

Fig. 1 **a** Dendrogram based on morphological traits constructed using Squared Euclidean distance and group average clustering method. **b** Biplot analysis based of PCA by morphological traits

in *Pongamia* based on pod and seed traits. Consequently, variations between these accessions are due to genotype-environment interactions and/or spontaneous mutations (Rajarajan et al., 2021). The clustering results suggest that these diverse genotypes are important in parental selection for tree improvement programs because they will offer unique alleles to have transgressive segregants for yield improvement (Rajarajan et al., 2021; Uchoi et al., 2016). Also, the biplot analysis in this study agreed with the clustering pattern based on morphological traits, clearly categorizing the accessions into two distinct coordinates as PC1 and PC2, with a total variability of 70.92% (Fig. 1b). This type of multivariate analysis for grouping accessions could provide tree breeders with an idea about accessions' variabilities.

SSR marker analysis for genetic diversity

In this study, out of fifteen SSR markers, eleven were found to be polymorphic, which revealed a high polymorphism across the eighteen accessions (Table 6). A total of 313 bands were produced by 11 primers; the alleles generated ranged from 2 to 6 with a mean of 3.6 per locus (Table 6). The primer P011 produced the maximum alleles (6) per locus (Fig. 2). The polymorphism percentage ranged from 50.0% to 100%, with a mean of 88.8%. Also, the PIC value

Table 6 Amplified markers, number of PCR amplified fragments and polymorphic information content obtained in SSR analysis of *Pongamia pinnata* accessions

Primers	Total alleles	Polymorphic alleles	Polymorphism %	Number of amplified fragments	PIC value
P001	4	4	100.00	28	0.68
P002	2	2	100.00	13	0.14
P003	4	2	50.00	70	0.75
P005	2	2	100.00	10	0.48
P006	3	3	100.00	30	0.53
P007	5	4	80.00	54	0.77
P008	5	5	100.00	14	0.55
P009	2	2	100.00	19	0.19
P010	3	3	100.00	28	0.54
P011	6	5	83.33	47	0.69
Total	36	32	–	313	–
Average	3.6	3.2	88.88	31.3	0.53

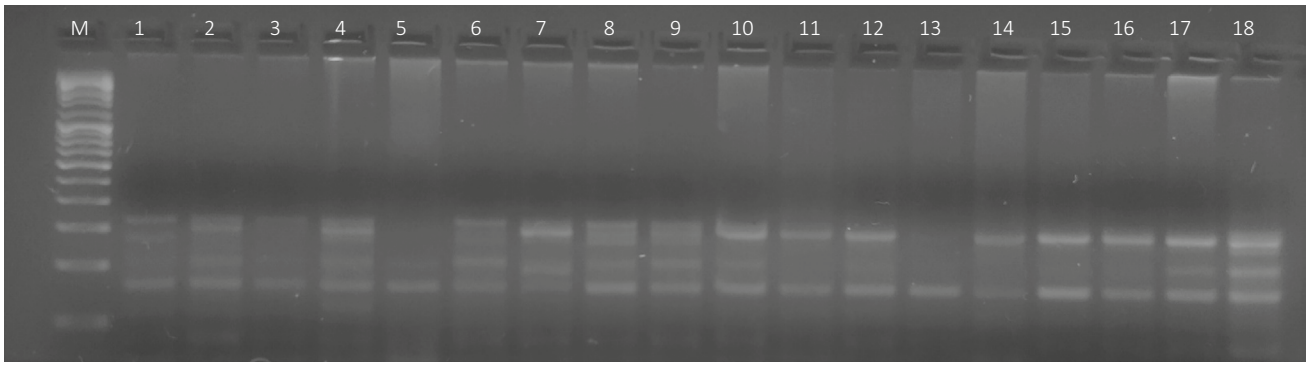

Fig. 2 SSR banding profile (P007) of 18 Pongamia accessions. Where M is 100 bp Ladder and 1–18 Pongamia accessions

Table 7 Analysis of molecular variance of 18 *Pongamia pinnata* accessions based on 11 SSR markers

Source	df	SS	MS	Est. Var	%
Among Pops	2	10.606	5.303	0.142	3%
Within Pops	15	69.227	4.615	4.615	97%
Total	17	79.833		4.757	100%
Nm	12.25**				

Nm = [(1/PhiPT) − 1]/4

PhiPT = AP/(WP + AP) = AP/TOT

AP = estimated variance among populations, WP = estimated variance within population

** denotes *P* value < 0.01

estimated for 11 primers ranged from 0.14 to 0.77 with an average of 0.53. P007 had the highest PIC value. The polymorphic information content of these SSR primers are higher than the earlier reports by different markers RAPD (Ahlawat et al., 2016; Kesari et al., 2010), ISSR (Kesari et al., 2010) AFLP (Kesari et al., 2010; Pavithra et al., 2013; Sharma et al., 2016, 2017) in *Pongamia*, indicating the discrimination potential of these SSR markers in genetic diversity studies (Table 7).

The similarity coefficient based on SSR markers ranged from 0.36 to 0.87, with an average of 0.55 indicating the existence of greater genetic diversity across the accessions. Cluster analysis based on the neighbor-joining method grouped all 18 accessions into five clusters according to their similarities (Fig. 3a). Cluster I was composed of six accessions, Cluster II and III each had four, cluster IV had three and cluster V had a single accession. According to the dissimilarity coefficient, the accessions NRCP11, NRCP14, and NRCP21 were identified as the most distinct and would provide divergent alleles for breeding programs (Rohini et al., 2020). Also, the principal coordinate analysis of the 18 accessions clearly divided the accessions into four different coordinates, which once again confirms the dendrogram pattern (Fig. 3b).

Population structure analysis

Information on genetic differentiation among and within populations is important for determining the relationship between accessions based on allele frequency distribution (Kumar et al., 2020; Rohini et al., 2020). In this study, we attempted to investigate the population structure and genetic relationships among *Pongamia* accessions using SSR markers (Fig. 4a). The two subpopulations (SUP1 & SUP2) were estimated based on the highest delta K value (K = 2) (Fig. 4b). The SUP1 and SUP2 had 11 and 7 accessions, respectively. Similarly, Pavithra et al., (2014) reported *Pongamia* population structure using AFLP markers. In addition, the structure analysis chart (red and green colors) using STRUCTURE HARVESTER clearly defined two gene pools present in this population, which also shows intermixing between the accessions based on the 20% genetic background of other accessions (Fig. 4b). Accordingly, NRCP11, NRCP12, NRCP16, and NRCP26 were found to be intermixed accessions. Similarly, Rohini et al., (2020) reported intermixed accessions in *Citrus jambhiri* population. Thus, aid in identifying pure genotypes from intermixed ones in a genetic resource repository for efficient utilization and characterization (Rajarajan et al., 2021; Rohini et al., 2020). Also, in this study, the results of structure analysis correspond with principal component analysis (PCoA) and neighbor-joining tree.

Furthermore, the AMOVA analysis instigated in this study found a high degree of genetic diversity within subpopulations and a low level among the population at 97% and 3%, respectively (Fig. 4c). The reason for high variation within the population might be due to outcrossing or spontaneous mutations. Furthermore, one could expect a high gene flow in the population by the low level of variation

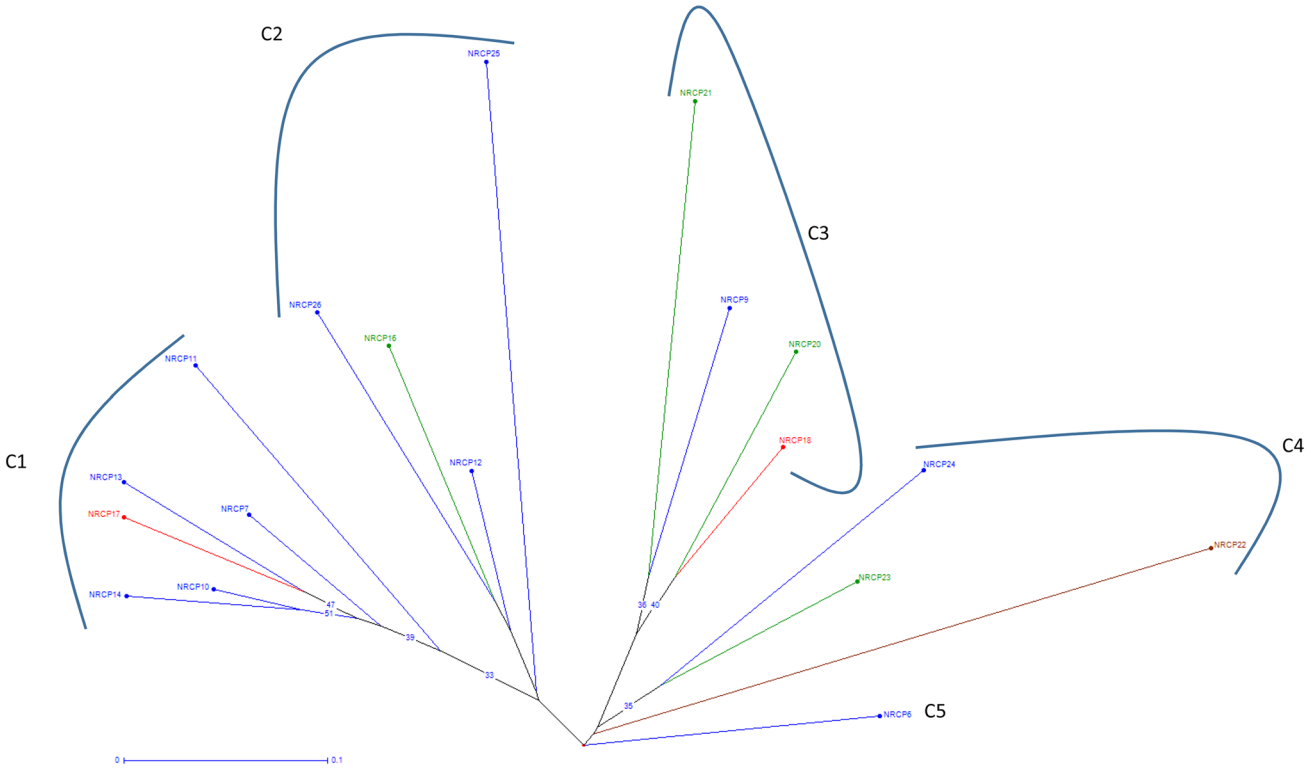

Fig. 3 **a** Radial neighbor-joining tree based on 74 alleles from 11 SSR loci among 18 Pongamia germplasm. **b** Principal coordinates analysis (PCoA) of 18 Pongamia accessions using SSR markers

among the population (Kumar et al., 2020). Likewise, in this study, a high level of gene flow justified only 3% differentiation among the population. This kind of study will help tree breeders to monitor and maintain genetic variation in the population for tree breeding programs.

Evaluation of morphological traits is a crucial step in determining genetic diversity and identifying superior genotypes (Rajarajan et al., 2021). Yet, the environment, which is considered a critical limitation, significantly impacts morphological trait expression. As a result, molecular markers should be used in concert with morphological traits to precisely measure the level of genetic diversity in the target species. Similarly, we assessed genetic diversity at both the morphological and molecular levels in this study. However, there was no association between the two assessments, which could be due to different phenotypes resulting from genotype and environment interaction. Overall, this study highlights the potential of combining markers and morphological traits to assess genetic differentiation and population structure in *Pongamia pinnata* for genetic improvement.

Conclusion

In this study, genetic variability for 100-seed weight, pod length, and the pod and seed thickness, had high heritability with high genetic advance, indicating that these traits are governed by additive genes and selection based on these traits would be very useful in trait improvement. Furthermore, multivariate analysis revealed that NRCP10 and NRCP20 are more diversified than other accessions. Also, these accessions have a high cetane index, and useful in

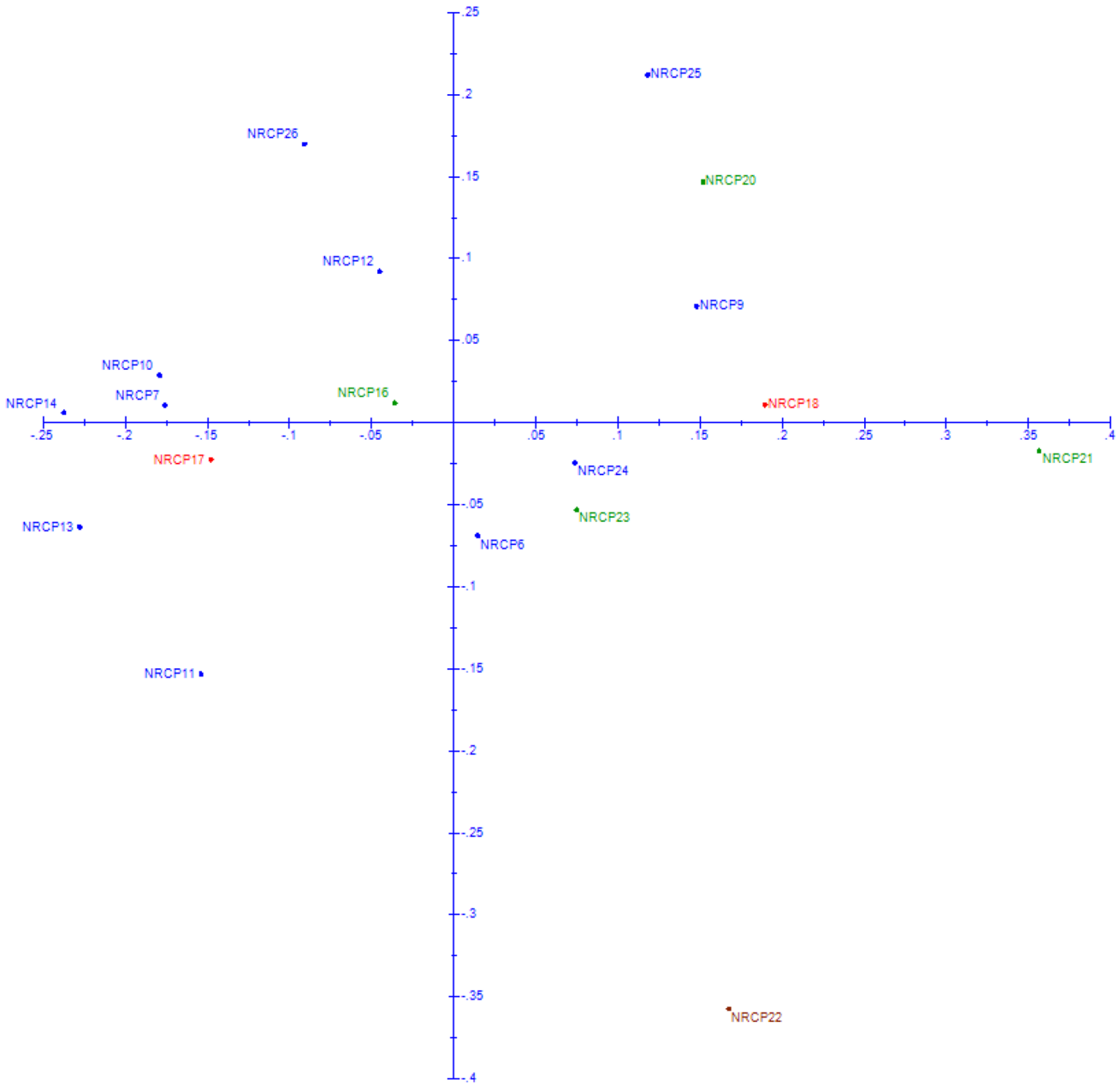

Fig. 3 (continued)

biodiesel production. In SSR marker analysis, accessions NRCP11, NRCP14, and NRCP21 found to be highly diversified. In addition, the population genetic structure analysis of *Pongamia* provided an idea about the population structure and its variability among and within the population for utilization in tree breeding.

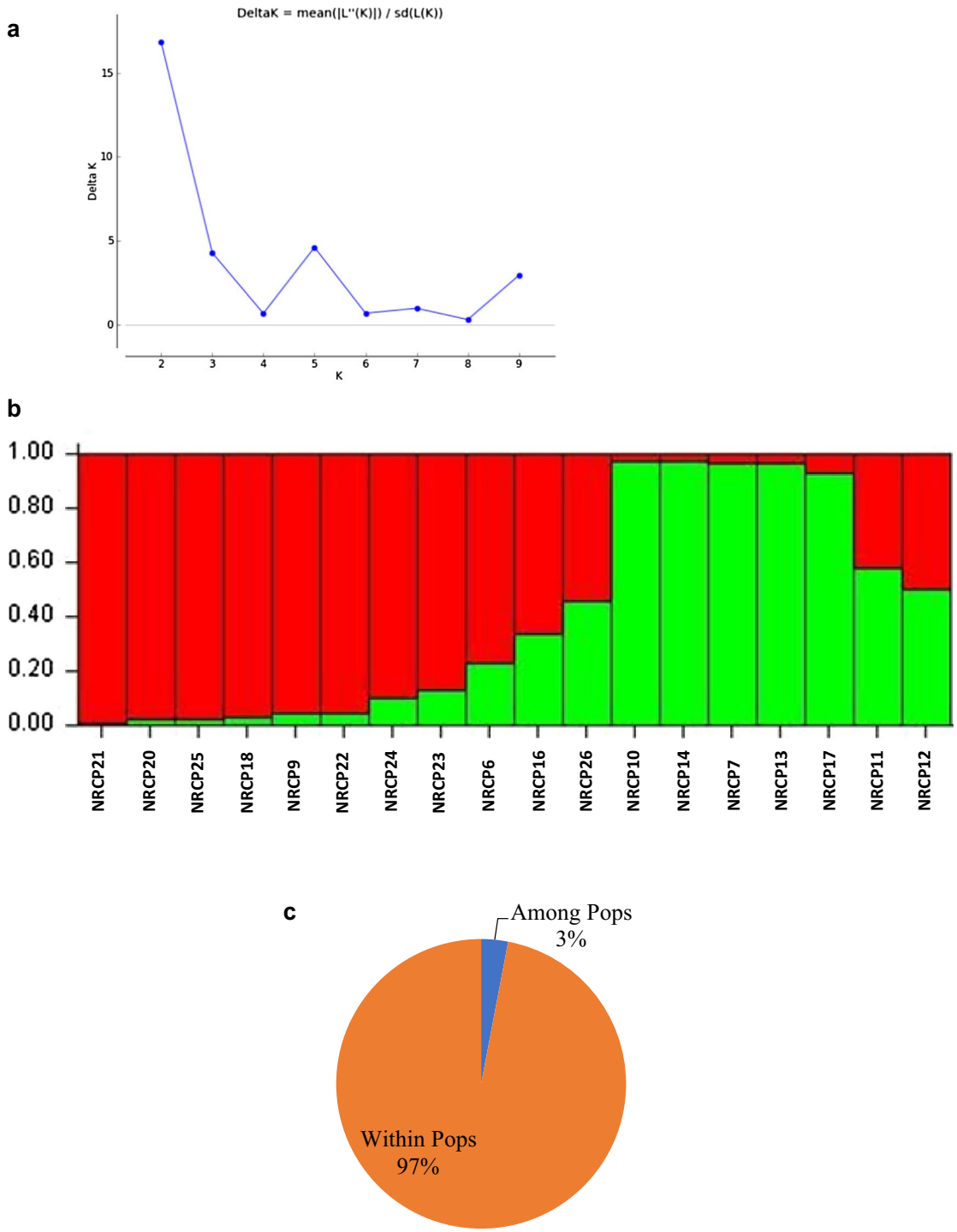

Fig. 4 Population structure analysis of *Pongamia pinnata* using 11 SSR markers: **a** Delta K analysis. **b** Structure analysis of two sub-populations of 18 Pongamia accessions and **c** analysis of molecular variance of 18 accessions of *Pongamia pinnata* based on SSR markers

Acknowledgements The authors are grateful to Field Assistant Mr. Rambabu, ICAR-Central Agroforestry Research Institute, India and Indian Council of Agricultural Research (ICAR), New Delhi, India, for facilities to complete this work.

Author contribution KR and AKH conceived the idea. KR wrote the main manuscript text. KR, AA, HA, MR, SS and RV prepared the manuscript. KR and AA revised the manuscript at different stages of the writing process and read and approved the revised manuscript.

Declarations

Conflict of interest All the authors declared that no conflicts of interest are associated with this publication.

References

Adjebeng-Danquah, J., Manu-Aduening, J., Asante, I. K., Agyare, R. Y., Gracen, V., & Offei, S. K. (2020). Genetic diversity and population structure analysis of Ghanaian and exotic cassava accessions using simple sequence repeat (SSR) markers. *Heliyon, 6*(1), e03154.

Ahlawat, S. P., Kumar, R. V., Ranjan, R., Pandey, S. K., Joshi, D. C., & Dhyani, S. K. (2016). Morphological and molecular level of genetic diversity among Pongamia [*Pongamia pinnata* (L.) Pierre] accessions. *Indian Journal of Biotechnology, 15*(1), 85–94.

Al Muqarrabun, L. M. R., Ahmat, N., Ruzaina, S. A. S., Ismail, N. H., & Sahidin, I. (2013). Medicinal uses, phytochemistry and pharmacology of *Pongamia pinnata* (L.) Pierre: A review. *Journal of Ethnopharmacology, 150*(2), 395–420.

Azam, M. M., Waris, A., & Nahar, N. M. (2005). Prospects and potential of fatty acid methyl esters of some non-traditional seed oils for use as biodiesel in India. *Biomass and Bioenergy, 29*(4), 293–302.

Badu-Apraku, B., Garcia-Oliveira, A. L., Petroli, C. D., Hearne, S., Adewale, S. A., & Gedil, M. (2021). Genetic diversity and population structure of early and extra-early maturing maize germplasm adapted to sub-Saharan Africa. *BMC Plant Biology, 21*(1), 1–15.

Botstein, D., White, R. L., Skolnick, M., & Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics, 32*(3), 314.

Doyle, J. J., & Doyle, J. L. (1990). Isolation ofplant DNA from fresh tissue. *Focus, 12*(13), 39–40.

Falush, D., Stephens, M., & Pritchard, J. K. (2007). Inference of population structure using multilocus genotype data: Dominant markers and null alleles. *Molecular Ecology Notes, 7*(4), 574–578.

Huang, J., Lu, X., Yan, H., Chen, S., Zhang, W., Huang, R., & Zheng, Y. (2012). Transcriptome characterization and sequencing-based identification of salt-responsive genes in *Millettia pinnata*, a semi-mangrove plant. *DNA Research, 19*(2), 195–207.

Irfan, M., Hayat, S., Hayat, Q., Afroz, S., & Ahmad, A. (2010). Physiological and biochemical changes in plants under waterlogging. *Protoplasma, 241*(1), 3–17.

Kalayasiri, P., Jeyashoke, N., & Krisnangkura, K. (1996). Survey of seed oils for use as diesel fuels. *Journal of the American Oil Chemists' Society, 73*(4), 471–474.

Karmee, S. K., & Chadha, A. (2005). Preparation of biodiesel from crude oil of *Pongamia pinnata*. *Bioresource Technology, 96*(13), 1425–1429.

Kaushik, N., Kumar, S., Kumar, K., Beniwal, R. S., Kaushik, N., & Roy, S. (2007). Genetic variability and association studies in pod and seed traits of *Pongamia pinnata* (L.) Pierre in Haryana, India. *Genetic Resources and Crop Evolution, 54*(8), 1827–1832.

Kesari, V., & Rangan, L. (2010). Development of *Pongamia pinnata* as an alternative biofuel crop—current status and scope of plantations in India. *Journal of Crop Science and Biotechnology, 13*(3), 127–137.

Kesari, V., & Rangan, L. (2011). Genetic diversity analysis by RAPD markers in candidate plus trees of *Pongamia pinnata*, a promising source of bioenergy. *Biomass and Bioenergy, 35*(7), 3123–3128.

Kesari, V., Madurai Sathyanarayana, V., Parida, A., & Rangan, L. (2010). Molecular marker-based characterization in candidate plus trees of *Pongamia pinnata*, a potential biodiesel legume. *AoB Plants, 2010*, 1–12.

Kesari, V., Krishnamachari, A., & Rangan, L. (2009). Effect of auxins on adventitious rooting from stem cuttings of candidate plus tree *Pongamia pinnata* (L.), a potential biodiesel plant. *Trees, 23*(3), 597–604.

Krisnangkura, K. (1986). A simple method for estimation of cetane index of vegetable oil methyl esters. *Journal of the American Oil Chemists' Society, 63*(4), 552–553.

Kumar, D., Chhokar, V., Sheoran, S., Singh, R., Sharma, P., Jaiswal, S., et al. (2020). Characterization of genetic diversity and population structure in wheat using array based SNP markers. *Molecular Biology Reports, 47*(1), 293–306.

Mukta, N., Murthy, I., & Sripal, P. (2009). Variability assessment in *Pongamia pinnata* (L.) Pierre germplasm for biodiesel traits. *Industrial Crops and Products, 29*(2–3), 536–540.

Naik, M., Meher, L., Naik, S., & Das, L. (2008). Production of biodiesel from high free fatty acid Karanja (*Pongamia pinnata*) oil. *Biomass and Bioenergy, 32*(4), 354–357.

Pavithra, H. R., Gowda, B., Prasanna, K. T., & Shivanna, M. B. (2013). Pod and seed traits in candidate plus trees of *Pongamia pinnata* (L.) Pierre from southern peninsular India in relation to provenance variation and genetic variability. *Journal of Crop Science and Biotechnology, 16*(2), 131–142.

Pavithra, H. R., Shivanna, M. B., Chandrika, K., Prasanna, K. T., & Gowda, B. (2014). Genetic analysis of *Pongamia pinnata* (L.) Pierre populations using AFLP markers. *Tree Genetics and Genomes, 10*(1), 173–188.

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research–an update. *Bioinformatics, 28*(19), 2537–2539.

Perrier, X., & Flori, A. (2003). Methods of data analysis. In *Genetic diversity of cultivated tropical plants* (pp. 47–80). CRC Press.

Rajarajan, K., Uthappa, A. R., Handa, A. K., Chavan, S. B., Vishnu, R., Shrivastava, A., et al. (2021). Genetic diversity and population structure of *Leucaena leucocephala* (Lam.) de Wit genotypes using molecular and morphological attributes. *Genetic Resources and Crop Evolution, 69*, 71–83.

Rajarajan, K., & Ganesamurthy, K. (2011). Genetic diversity analysis of sorghum [*Sorghum bicolor* (L.) Moench] genotypes for drought tolerance using SSR markers. *Indian Journal of Genetics and Plant Breeding, 71*(1), 17.

Rohini, M. R., Sankaran, M., Rajkumar, S., Prakash, K., Gaikwad, A., Chaudhury, R., & Malik, S. K. (2020). Morphological characterization and analysis of genetic diversity and population structure in *Citrus× jambhiri* Lush. using SSR markers. *Genetic Resources and Crop Evolution, 67*, 1259–1275.

Scott, P. T., Pregelj, L., Chen, N., Hadler, J. S., Djordjevic, M. A., & Gresshoff, P. M. (2008). *Pongamia pinnata*: An untapped resource for the biofuels industry of the future. *Bioenergy Research, 1*(1), 2–11.

Sharma, S. S., Islam, M. A., Malik, A. A., Kumar, K., Negi, M. S., & Tripathi, S. B. (2016). Seed traits, fatty acid profile and genetic diversity assessment in *Pongamia pinnata* (L.) Pierre germplasm. *Physiology and Molecular Biology of Plants, 22*(2), 193–205.

Sharma, S. S., Islam, M. A., Singh, V. K., Negi, M. S., & Tripathi, S. B. (2017). Genetic diversity, population structure and association study using TE-AFLP markers in *Pongamia pinnata* (L.) Pierre germplasm. *Tree Genetics & Genomes, 13*(1), 1–10.

Shivanna, M. B., & Rajakumar, N. (2010). Ethno-medico-botanical knowledge of rural folk in Bhadravathi taluk of Shimoga district, Karnataka. *Indian Journal of Traditional Knowledge, 9*(1), 158–162.

Uchoi, A., Malik, S. K., Choudhary, R., Kumar, S., Rohini, M. R., Pal, D., et al. (2016). Inferring phylogenetic relationships of Indian

citron (*Citrus medica* L.) based on rbcL and matK sequences of chloroplast DNA. *Biochemical Genetics, 54*(3), 249.
Yadav, P. P., Ahmad, G., & Maurya, R. (2004). Furanoflavonoids from *Pongamia pinnata* fruits. *Phytochemistry, 65*(4), 439–443.